

UNIVERSITÉ DU QUÉBEC À MONTRÉAL

L'ASSIGNATION DU GENRE GRAMMATICAL EN FRANÇAIS LANGUE SECONDE :
TRANSFERT OU TERMINAISONS DES NOMS?

MÉMOIRE
PRÉSENTÉ
COMME EXIGENCE PARTIELLE
DE LA MAÎTRISE EN LINGUISTIQUE

PAR
FRÉDÉRIC GAREAU

JUIN 2008

UNIVERSITÉ DU QUÉBEC À MONTRÉAL
Service des bibliothèques

Avertissement

La diffusion de ce mémoire se fait dans le respect des droits de son auteur, qui a signé le formulaire *Autorisation de reproduire et de diffuser un travail de recherche de cycles supérieurs* (SDU-522 – Rév.01-2006). Cette autorisation stipule que «conformément à l'article 11 du Règlement no 8 des études de cycles supérieurs, [l'auteur] concède à l'Université du Québec à Montréal une licence non exclusive d'utilisation et de publication de la totalité ou d'une partie importante de [son] travail de recherche pour des fins pédagogiques et non commerciales. Plus précisément, [l'auteur] autorise l'Université du Québec à Montréal à reproduire, diffuser, prêter, distribuer ou vendre des copies de [son] travail de recherche à des fins non commerciales sur quelque support que ce soit, y compris l'Internet. Cette licence et cette autorisation n'entraînent pas une renonciation de [la] part [de l'auteur] à [ses] droits moraux ni à [ses] droits de propriété intellectuelle. Sauf entente contraire, [l'auteur] conserve la liberté de diffuser et de commercialiser ou non ce travail dont [il] possède un exemplaire.»

REMERCIEMENTS

Ce travail n'aurait pu être effectué sans l'heureuse participation de plusieurs personnes. Nous nous devons de remercier tout d'abord monsieur Robert Papen, directeur de recherche pour ce mémoire, de qui nous avons reçu une aide constante et éclairante. L'expérience, sa rigueur et sa passion pour le domaine font de lui un excellent chercheur et un excellent guide. Nous remercions aussi tout particulièrement madame Marie Labelle et monsieur Thomas Cobb pour avoir accepté de faire partie de notre comité de lecture.

Sans l'aide de monsieur Bertrand Fournier, coordonnateur et analyste statisticien pour le Service de consultation en analyse de données du département de mathématiques de l'UQAM (SCAD), nous n'aurions pas pu générer les données statistiques de nos tests. Michel Bastien, concepteur du logiciel LÉA, nous a donné patiemment et à multiples reprises ses conseils et son aide technique pour la passation des tests sur ordinateur. Maria Fernanda Zaldívar Turrent et Francisco Goire nous ont donné leurs précieux conseils linguistiques en tant qu'hispanophones. Claude Dionne, Ève Langevin, Ghislaine Gagné, John Lumsden, Luc Thériault, Sophia Stevenson et Xavier Cornillier nous ont conseillé à plusieurs reprises. Catherine Duranleau, Christophe Tessier, Claude Timmons et Juliane Bertrand nous ont fait accéder à leurs classes d'étudiants de l'École de langues. Sans cette aide, il aurait été difficile de constituer une banque de participants pour cette recherche. Au total, ce sont 60 hispanophones et sinophones qui ont bien voulu nous donner chacun au moins une demi-heure de leur temps. Sans cette collaboration, ce projet n'aurait pas pu être mené à terme.

Tous les gens mentionnés ci-dessus sont donc chaleureusement remerciés. Nous remercions aussi toutes les personnes que nous aurions pu oublier.

TABLE DES MATIÈRES

LISTE DES FIGURES.....	vi
LISTE DES TABLEAUX	vii
LISTE DES ABRÉVIATIONS, SIGLES ET ACRONYMES	ix
RÉSUMÉ.....	xi
INTRODUCTION	1
CHAPITRE I	
PROBLÉMATIQUE ET CADRE THÉORIQUE.....	3
1.1 Introduction	3
1.2 La problématique	3
1.3 Le genre grammatical.....	6
1.3.1 Le classement du lexique.....	6
1.4 Le français et le genre grammatical	10
1.4.1 Quelques distinctions.....	10
1.4.2 Les terminaisons des noms	12
1.5 Le transfert	13
1.5.1 Le transfert en général	14
1.5.2 Le transfert et le genre grammatical	17
1.5.3 Les congénères et le transfert.....	18
1.6 Deux études importantes	20
1.6.1 Sabourin, Stowe et De Haan (2006).....	21
1.6.2 Lyster (2006).....	23
1.7 L'espagnol et le chinois.....	27
1.7.1 L'espagnol	27
1.7.2 Le chinois.....	29
1.8 Les questions de recherche et les hypothèses	29

1.8.1	Questions de recherche et transfert.....	30
1.8.2	Questions de recherche et terminaisons.....	32
1.8.3	Autres questions de recherche.....	32
1.9	Conclusion.....	34
CHAPITRE II		
MÉTHODOLOGIE.....		35
2.1	Introduction	35
2.2	Les participants.....	36
2.3	Le corpus de Baudot (1992)	38
2.4	La confection des tests	38
2.5	La passation des tests	46
2.6	Les procédés de compilation et d'analyse des résultats	48
2.7	Conclusion.....	49
CHAPITRE III		
PRÉSENTATION ET ANALYSE DES RÉSULTATS.....		50
3.1	Introduction	50
3.2	La réponse à la question de recherche générale.....	51
3.3	Les réponses aux questions de recherche sur le transfert	51
3.3.1	Les hispanophones	52
3.3.2	Les congénères.....	52
3.4	Les réponses aux questions de recherche sur les terminaisons.....	55
3.4.1	Les terminaisons marquées pour le genre.....	56
3.4.2	Les terminaisons en -e muet.....	58
3.5	Les réponses aux autres questions de recherche	61
3.5.1	La fréquence des noms	63
3.5.2	Le niveau de connaissance du vocabulaire des participants	66
3.5.3	L'hypothèse du masculin par défaut.....	68
3.5.4	La variable « sexe ».....	70
3.6	Les non-mots	70
3.7	Les non-réponses	73
3.8	Les résultats marginaux.....	74
3.9	Conclusion.....	74

CHAPITRE IV	
DISCUSSION GÉNÉRALE.....	76
4.1 Introduction	76
4.2 Les liens avec les deux études importantes	76
4.2.1 L'effet du transfert sur l'assignation du genre	76
4.2.2 L'effet des terminaisons sur l'assignation du genre	79
4.3 Une observation.....	80
4.4 Les limites et les recommandations	80
4.5 Conclusion.....	83
CHAPITRE V	
IMPLICATIONS PÉDAGOGIQUES.....	84
5.1 Notre modèle général d'acquisition du genre grammatical	84
5.2 Les considérations générales	85
5.3 Nos suggestions d'applications pédagogiques.....	87
5.4 Conclusion.....	89
CONCLUSION.....	90
APPENDICE A	
DOCUMENTS UTILISÉS.....	92
RÉFÉRENCES	100

LISTE DES FIGURES

Figure	page
1.1	Résultats pour l'expérience I (tiré de Sabourin, Stowe et de Haan, 2006, p. 11)..... 23

LISTE DES TABLEAUX

Tableau	page
1.1 Exemples de classifieurs (compteurs) en japonais (tiré et traduit de Wikipedia).....	7
1.2 Exemples de classeurs nominaux en swahili (tiré et traduit de Hudson, 2000, p. 77).....	8
1.3 Moyennes des résultats (et écarts) pour chaque groupe langagier, par type d'item (tiré et traduit de Sabourin, Stowe et de Haan, 2006, p. 10)....	22
1.4 Terminaisons typiquement féminines (tiré de Lyster, 2006, p. 87).....	25
1.5 Terminaisons typiquement masculines (tiré de Lyster, 2006, p. 87).....	25
1.6 Terminaisons typiquement ambiguës (tiré de Lyster, 2006, p. 88).....	26
2.1 Les 15 non-mots.....	41
2.2 La répartition des terminaisons des 60 noms (vrais mots).....	44
2.3 Les 60 noms choisis (vrais mots) avec leur fréquence dans le corpus de Baudot.....	45
3.1 Pourcentage moyen de réussite (avec écart type) au test d'assignation du genre selon « même genre » et « congénère ».....	52
3.2 Comparaison des pourcentages moyens de réussite des deux groupes linguistiques, avec le test bilatéral de Wilcoxon utilisant l'approximation normale, selon la combinaison des catégories « même genre » et « congénère » (test d'assignation du genre).....	54

3.3	Pourcentage moyen de réussite (avec écart type) au test d'assignation du genre selon « genre du nom » et « terminaison marquée »	56
3.4	Comparaison des pourcentages moyens de réussite des deux groupes linguistiques, avec le test bilatéral de Wilcoxon utilisant l'approximation normale, selon la combinaison des catégories « genre du nom » et « terminaison marquée » (test d'assignation du genre).....	57
3.5	Pourcentage moyen de réussite (avec écart type) au test d'assignation du genre selon « genre du nom » et « terminaison en -e »	59
3.6	Comparaison des pourcentages moyens de réussite des deux groupes linguistiques, avec le test de Wilcoxon bilatéral utilisant l'approximation normale, selon la combinaison des catégories « genre du nom » et « terminaison en -e » (test d'assignation du genre)	60
3.7	Pourcentage moyen de réussite (avec écart type) au test de reconnaissance selon « même genre » et « congénère »	61
3.8	Comparaison des pourcentages moyens de réussite des deux groupes linguistiques, avec le test bilatéral de Wilcoxon utilisant l'approximation normale, selon la combinaison des catégories « même genre » et « congénère » (test de reconnaissance)	63
3.9	Distribution des erreurs selon le genre grammatical pour le groupe des hispanophones	68
3.10	Distribution des erreurs selon le genre grammatical pour le groupe des sinophones	69
3.11	Pourcentage de réussite au test d'assignation du genre, selon les terminaisons en -e muet final, pour les dix non-mots à terminaisons marquées pour le genre.....	71

LISTE DES ABRÉVIATIONS, SIGLES ET ACRONYMES

C	Congénères (du français à l'espagnol)
e	Terminaison en -e
F (parfois f)	Féminin
H	Hispanophones
L1	Langue première
L2	Langue seconde
m	Terminaison marquée pour le genre
M (parfois m)	Masculin
MF	Masculin / Féminin (test d'assignation du genre)
MG	Même genre (en français et en espagnol)
N	Non
NC	Non congénères (du français à l'espagnol)
Ne	Non (pas de) terminaison en -e
NIV	Nom à initiale vocalique
Nm	Terminaison non marquée pour le genre

NMG	Non (pas le) même genre (en français et en espagnol)
O	Oui
ON	Oui / Non (test de reconnaissance)
Q	Question
RAEi	Real Academia Española informatisé
S	Sinophones
SCAD	Service de consultation en analyse de données
TL	Target language (langue cible)
TLFi	Trésor de la Langue Française informatisé

RÉSUMÉ

Ce mémoire traite de l'assignation du genre grammatical, élément particulièrement difficile à maîtriser par les apprenants du français langue seconde. Il vise à déterminer si le transfert de la langue première (L1) est un facteur majeur (transfert positif ou négatif) pour cette question et à comparer cette stratégie potentielle à celle du recours aux terminaisons des noms. Plus spécifiquement, une comparaison a été faite entre les résultats d'assignation du genre pour une variété de substantifs français par des hispanophones et des sinophones (mandarin) apprenant le français. Une telle étude est d'un grand intérêt puisque l'espagnol possède deux genres, comme en français, mais distribués différemment, alors que le mandarin n'a pas de genre. Les participants étaient tous des adultes. Leur niveau de connaissance du vocabulaire de nos tests a été mesuré et il a déterminé, entre autres, le classement des participants. Il a ensuite été pris en considération lors de l'analyse des résultats. D'autres variables ont été observées (par ex. : l'influence du sexe des participants sur l'assignation du genre grammatical).

Pour répondre à nos questions de recherche, les participants (60) ont été conviés à passer deux tests individuels, présentés sur support informatique : un de reconnaissance des noms et un d'assignation du genre grammatical. Les principales questions de recherche étaient de savoir quel serait le groupe qui obtiendrait les meilleurs résultats au test d'assignation du genre grammatical et concernaient le transfert et les terminaisons des noms.

Les deux groupes linguistiques ont obtenu des scores globaux d'assignation du genre similaires, mais avec des résultats distribués très différemment. Nous avons confirmé nos hypothèses principales : les hispanophones ont obtenu des résultats reflétant systématiquement un transfert de leur L1, et ce, surtout lorsque les noms observés étaient congénères avec leurs homologues en espagnol. Si les congénères étaient du même genre, l'assignation correcte était de 96,0 %, donc un effet de transfert positif très évident. Par contre, si les congénères étaient d'un genre différent, l'assignation correcte du genre français était de 33,7 %, donc un effet évident de transfert négatif (un haut taux d'erreurs). Pour les sinophones, les résultats ont coïncidé avec une attribution du genre masculin lorsque les noms français ne possédaient pas de *-e* muet final et du genre féminin lorsque les noms français possédaient un *-e* muet final. Nous en avons conclu des retombées pédagogiques qui favorisent un enseignement différent du genre grammatical selon le groupe linguistique.

Mots-clés : acquisition, chinois, congénères, didactique, espagnol, français, genre grammatical, langue seconde, lexique, linguistique, terminaisons, transfert

INTRODUCTION

Un des problèmes majeurs qu'ont les apprenants du français langue seconde est la question de l'assignation correcte du genre. Ce problème ne semble s'amenuiser ni avec le temps ni avec le niveau de compétence acquise dans la langue. Tucker, Lambert et Rigault (1977, p. 11) considèrent que « [F]rom a practical point of view, the necessity to master gender may be the single most frustrating and difficult part of the study of French as a second language, both for those whose native language lacks gender distinctions, and for those whose language makes use of some other basis for such distinction. » Par contre, Franceschina (2005, p. 114) souligne : « Children acquiring languages with grammatical gender as their L1 show remarkable speed and few developmental errors compared to adult L2 learners. » Il semblerait donc que les difficultés que pose l'assignation du genre soient limitées aux apprenants adultes (n'excluant pas les adolescents) de langue seconde.

Parmi les facteurs linguistiques impliqués, l'influence du transfert de la langue première (L1) ainsi que celle des terminaisons des noms sont deux grandes hypothèses concernant ce qui nous intéresse. Pour cette dernière, Tucker, Lambert et Rigault (1977) ont démontré que la terminaison des noms était un puissant indice de genre, et que les locuteurs natifs du français assignaient très bien le genre à des non-mots ayant différentes terminaisons typiquement féminines ou typiquement masculines. En ce qui concerne le transfert et l'assignation du genre, peu a été dit, comme nous le verrons.

Nous proposons, dans ce mémoire, de déterminer le rôle du recours au transfert lors de l'assignation du genre grammatical en français langue seconde. Ce rôle sera notamment comparé à celui du recours aux terminaisons des noms. Pour ce faire, nous nous appuierons principalement sur les recherches de Sabourin, Stowe et de Haan (2006), pour ce qui est du transfert, et celles de Lyster (2006), en ce qui concerne les terminaisons des noms.

Ce mémoire est constitué de cinq chapitres. Dans le premier, Problématique et cadre théorique, nous décrirons la question d'ensemble, bien appuyée sur les écrits pertinents, pour ensuite poser nos questions de recherche. Le deuxième chapitre, Méthodologie, fera place à la description de notre processus de conception des tests, qui nous permettront de répondre à nos questions de recherche et de confirmer ou d'infirmer nos hypothèses. Le troisième chapitre, Présentation et analyse des résultats, brossera un tableau des statistiques obtenues et de l'interprétation des données qui en a résulté. Ensuite, le chapitre quatre, Discussion générale, fera les liens nécessaires entre les résultats obtenus et les écrits consultés. Il permettra aussi de parler des limites du mémoire et de faire un retour critique sur celui-ci tout en ouvrant sur les recherches à venir. Le cinquième et dernier chapitre, Implications pédagogiques, proposera des retombées pédagogiques des résultats du mémoire. À la fin de l'ouvrage, afin de permettre au lecteur de bien visualiser les outils utilisés lors de la passation des tests, certains documents pertinents seront annexés.

CHAPITRE I

PROBLÉMATIQUE ET CADRE THÉORIQUE

1.1 Introduction

Dans ce chapitre, nous formulerons d'abord la problématique de la recherche. Ensuite, nous spécifierons et définirons les différents termes nécessaires à la bonne compréhension du mémoire, tout en nous appuyant sur les différents écrits existant sur le sujet. Nous décrirons aussi en détail les deux articles sur lesquels notre recherche s'appuie. En dernier lieu, nous poserons nos questions de recherche et émettrons certaines hypothèses.

1.2 La problématique

Le genre grammatical est, lors de l'apprentissage du français par des non-francophones, un des aspects les plus difficiles à acquérir convenablement sur le long terme. Même à niveau très avancé, les apprenants ont encore tendance à commettre des erreurs de genre. Voilà donc pourquoi il est nécessaire d'étudier ce phénomène en détail, à savoir s'il serait possible de remédier à cette lacune de l'acquisition du français langue seconde. Notons ici que cette lacune en est une concernant la précision langagière, et non l'aisance à communiquer. En effet, la mauvaise utilisation du genre grammatical pour un nom donné n'empêchera pas nécessairement le nouvel utilisateur de la langue française de communiquer. Cependant, cela pourra parfois occasionner quelques ambiguïtés sémantiques, notamment lors de l'accord en genre ou lorsqu'il y a des homophones distingués par des genres différents.

En français langue première, le genre grammatical semble s'acquérir de façon naturelle et inconsciente, et ne semble pas poser de problèmes majeurs aux enfants. Mais en est-il de même pour le français langue seconde chez les adultes? En effet, « a series of studies was prompted by observations that native speakers make gender assignments correctly even

though they do not know “why” or “how” they make these choices, and that non-natives, even those who have studied French for many years, are consistently poor and confused in deciding about gender » (Tucker, Lambert et Rigault, 1977, p. 63). Il y a donc de fortes raisons de croire que l'apprentissage du genre grammatical dans une langue seconde (du moins pour ce qui est du français) ne se produit pas de la même façon que dans une langue première. Pour les locuteurs natifs, il semble que ce soit un processus inexplicable : « Although the native speakers are highly skilled at making gender assignments, they are unable to explain this ability » (Tucker, Lambert et Rigault, 1977, p. 64). Les apprenants du français langue seconde ne semblent pas acquérir le genre grammatical de façon naturelle et inconsciente puisque ce serait « an area of the grammar known to be problematic for L2 learners » (Franceschina, 2005, p. 6). Ils ont besoin de stratégies spécifiques, et celles-ci sont peut-être différentes d'un groupe linguistique (d'une L1) à l'autre. C'est là notre hypothèse provisoire, du moins.

Comme nous l'avons déjà mentionné dans l'introduction, notre étude cherche à savoir si la stratégie du recours au transfert est efficace ou non lors de l'assignation du genre grammatical en français langue seconde. Lado (1957, p. 2) considère que : « [...] individuals tend to transfer the forms and meanings, and the distribution of forms and meanings of their native language and culture to the foreign language and culture [...]. » De plus, cette stratégie cherche à être comparée à d'autres stratégies possibles. Celle du recours aux terminaisons des noms est la plus connue et la plus documentée. Nous essayerons de voir plus clair dans l'ensemble des stratégies d'apprentissage possibles, mais en gardant toujours bien en vue ces deux dernières, les principales.

Lorsqu'on fait une recherche de ce type, il est important, d'une part, de prévoir le plus possible les stratégies qui pourraient être utilisées par les participants. Il est, d'autre part, impératif de bien isoler chacun des facteurs qui pourraient éventuellement influencer sur les résultats. Une vigilance particulière doit donc être apportée aux différents aspects de la langue française, mais aussi de la langue des participants à l'étude. Dans cette recherche, nous aurons recours à des locuteurs de l'espagnol (hispanophones) et à des locuteurs de langues chinoises (sinophones), principalement le mandarin. Il sera donc nécessaire de bien connaître les caractéristiques principales du lexique nominal de ces deux langues. Ce lexique

est-il similaire, du point de vue de la forme, à celui du français? Possède-t-il les mêmes caractéristiques intrinsèques, le même comportement grammatical, en relation aux autres mots?

Nous avons choisi deux langues au comportement lexical fort différent. L'espagnol comporte, dans l'ensemble, les mêmes caractéristiques que le français. Il possède deux genres, le masculin et le féminin. Dans cette langue, le genre est aussi une caractéristique du nom. Plusieurs catégories de mots, comme l'adjectif, s'accordent avec le nom, notamment en genre. De plus, le vocabulaire espagnol est souvent similaire, du point de vue de la forme, à celui du français. Il comporte ce qu'on appelle des noms « congénères »¹ avec ceux du français (une planète (fr.) / *un planeta* (esp.)). Cependant, ces noms espagnols, congénères ou non, n'ont pas toujours le même genre qu'en français (planète = fém. / *planeta* = masc.). Une question se pose alors : les apprenants hispanophones du français langue seconde ont-ils de la facilité ou non à assigner correctement le genre à des noms français ou à accorder en genre ces mêmes noms, étant donné qu'ils ont déjà la connaissance des mécanismes du genre grammatical, qu'ils possèdent les mêmes genres qu'en français (masculin et féminin) et qu'ils peuvent reconnaître assez souvent les noms qu'ils lisent ou entendent dans cette langue seconde, surtout si ce sont des congénères?

Pour pouvoir bien répondre à ces questions, il est important de choisir un autre groupe linguistique qui ne possède aucunement les caractéristiques de l'espagnol ou du français, et qui pourrait servir de groupe contrôle. Pour ce faire, nous avons choisi les sinophones. Lorsqu'on parle de chinois, on doit parler « des » langues chinoises. Le mandarin est la langue chinoise la plus utilisée, mais des locuteurs de plusieurs langues chinoises sont regroupés sous l'appellation « sinophones ». Un sinophone est donc quelqu'un qui parle une langue chinoise parmi tant d'autres. La majorité des participants sinophones à cette recherche parlent mandarin, sauf quelques rares exceptions. Nous emploierons dorénavant le mot « chinois » pour toutes les langues chinoises puisque celles-ci ont les mêmes caractéristiques en ce qui concerne ce qui nous intéresse ici, c'est-à-dire leur comportement lexical en comparaison au français.

La langue chinoise est donc idéale pour bien représenter les caractéristiques opposées à l'espagnol en ce qui concerne le lexique. Le chinois ne possède pas de genre grammatical. Il a une autre façon de classer le lexique, qui sera discutée ultérieurement. De plus, aucun nom n'est congénère avec ceux du français (ou, en l'occurrence, de l'espagnol). Il est donc absolument impossible de transférer, à partir du chinois, une caractéristique lexicale qui pourrait aider à assigner le genre aux noms français. Cependant, le recours aux terminaisons ou à d'autres stratégies d'assignation du genre grammatical peut être vérifié plus facilement qu'avec les hispanophones, puisque les variables reliées au transfert lexical (le même genre grammatical ou non, la congénarité des noms ou non) sont évitées, ce qui aide à contrôler les données.

C'est donc à partir de cette problématique et de ces écueils potentiels que nous allons discuter le cadre théorique de notre recherche en détail, et décrire les quelques recherches pertinentes qui ont été publiées.

1.3 Le genre grammatical

Nous nous permettons de débiter cette section par deux citations essentielles: « The category of gender remains a puzzle to many linguists who do not agree on the origin or even on a definition of gender » (Ibrahim, 1973, p. 11) et « Gender is a fascinating category, central and pervasive in some languages and totally absent in others » (Corbett, 1991, pages liminaires). À la lumière de ces citations, il est primordial de tenter de bien définir ce qu'est le genre grammatical, catégorie difficile à définir. En général, le genre est une façon de classer les noms. En effet, « gender is not essentially different from other systems of noun classification which are known to exist in some language groups such as Bantu » (Ibrahim, 1973, p. 11). Nous allons donc comparer le genre à d'autres façons de classer le lexique.

1.3.1 Le classement du lexique

Les langues du monde ont, en effet, différentes façons de classer le lexique. Trois groupements peuvent être faits en ce qui concerne le classement des noms, bien que certaines

¹ Voir la section 1.5.3.1 pour la définition.

langues ne soient pas associées à ceux-ci (par ex. l'anglais). Une première classification se retrouve principalement dans les langues asiatiques (les classificateurs). Un autre type de classement est plus souvent observé dans les langues africaines (les classes nominales). Finalement, le dernier groupe de langues est le plus nombreux et comporte la majeure partie des langues indo-européennes : le genre grammatical.

1.3.1.1 Les classificateurs

Dans certaines langues asiatiques, on retrouve un type de classification du lexique basé sur des compteurs. Parfois appelées classificateurs (ou classifieurs), ces particules sémantiques permettent de classer les noms lorsque ceux-ci sont présentés avec un nombre.

Tableau 1.1
Exemples de classifieurs (compteurs) en japonais
(tiré et traduit de Wikipedia)

<u>Japonais</u>	<u>Transcription occidentale (romaji)</u>	<u>Traduction littérale française</u>	<u>Français</u>
鉛筆五本	enpitsu go- <u>hon</u>	crayon-cinq- <u>choses cylindriques</u>	cinq crayons
犬三匹	inu san- <u>biki</u>	chien-trois- <u>choses animales</u>	trois chiens
子供四人	kodomo yo- <u>nin</u>	enfant-quatre- <u>choses humaines</u>	quatre enfants

On peut retrouver, dans une langue comme le mandarin ou le japonais, une ou deux centaines de classificateurs. Par exemple, « in languages with such a system, a numeral (e.g. 'three') cannot be combined with a noun (e.g. 'book') directly, but requires the additional use of a classifier » (Hellinger et Bussmann, 2001, p. 4). De cette façon, pour exprimer l'équivalent du syntagme nominal français *trois livres*, il faut insérer un classificateur, ce qui donne *trois / objet plat / livres* (Greenberg, 1972, p. 5, traduction libre de l'anglais). Dans cet exemple, la particule exprimée par l'expression *objet plat* représente le classificateur, qui classe un livre dans la catégorie des objets plats.

In Japanese grammar, most nouns are effectively mass nouns, and measure words must be used with a number when counting them. The appropriate measure word is chosen based on the kind and shape of the noun, and combines with the numeral, sometimes adopting several different forms. This is similar to noun classes in many African languages, except that the classifiers are used only when counting.

Cette citation, tirée du Wikipedia anglais, prépare bien le lecteur à la compréhension des quelques exemples du tableau 1.1.

1.3.1.2 Les classes nominales

Dans certaines langues bantoues, comme le swahili, il existe une autre façon d'ordonner le lexique (Hellinger et Bussmann, 2001, p. 5). On parle ici d'une forme de classeurs sémantiques beaucoup plus restreinte. Ces classes nominales, catégories morpho-sémantiques permettant de classer le lexique, ont un nombre dépassant rarement 20, et se rapprochent du genre grammatical que l'on retrouve dans les langues indo-européennes. Cependant, les liens sémantiques de ces classes nominales sont encore assez concrets, en ce sens que ces particules trient les noms d'une façon qui représente plutôt bien la réalité des référents.

Tableau 1.2
Exemples de classeurs nominaux en swahili
(tiré et traduit de Hudson, 2000, p. 77)

<u>Préfixe</u>	<u>Description de la classe</u>	<u>Exemple</u>	<u>Glose</u>
<u>wa-</u>	choses animées	<u>w</u>avulana	'garçons'
<u>mi-</u>	plantes	<u>m</u>iti	'arbres'
<u>vi-</u>	instruments	<u>v</u>iti	'lits'

Certains voient ces classeurs sémantiques (habituellement sous forme de préfixes) comme une forme de genre grammatical. « Typical Bantu languages have several genders reflected in complex agreement systems » (Corbett, 1991, p. 43). Dans le tableau 1.2, tiré et traduit de Hudson (2000), il y a trois exemples de noms pluriels comportant des classeurs nominaux comme préfixes.

1.3.1.3 Genre grammatical

Dans les langues indo-européennes, le nombre de catégories de classement se restreint à deux ou trois par langue. On parle ici véritablement de « genre grammatical » (tous les chercheurs s'entendent). Mais si tous les chercheurs ne s'entendent pas nécessairement sur la nature même du genre grammatical, la plupart le mettent en relation avec les différents classeurs lexicaux, tels que dans les langues de l'Asie et de l'Afrique. En effet, « [...] on a sample of 174 languages, just over one-fourth of them had gender or some other form of nominal classification system » (Franceschina, 2005, p. 69).

Les genres qu'on retrouve habituellement sont le masculin, le féminin, et parfois le neutre. En latin, en allemand ou en russe, les trois genres sont présents. Des langues comme le français, l'espagnol ou l'italien ont perdu le genre neutre du latin. D'autres langues comme le néerlandais ou les langues scandinaves possèdent d'autres genres, souvent incluant le neutre, puis une autre forme, qui remplace le masculin et le féminin. Le genre grammatical est donc le plus souvent binaire ou ternaire dans une langue donnée. De plus, cette façon de classer le lexique, tout en aidant la bonne cohérence grammaticale par l'accord, a perdu en grande partie ses propriétés sémantiques (ou référentielles). En effet, la plupart des noms inanimés ont un genre qui n'a aucun rapport avec une classe sémantique comme dans les langues asiatiques ou africaines. Nous parlons donc ici d'un moyen non sémantique, et en apparence arbitraire, de classer les noms d'une langue. Hellinger et Bussmann (2001, p. 3) donnent l'exemple du mot *soleil*, qui possède un genre différent selon la langue, ce qui confirme l'aspect sémantiquement arbitraire du genre grammatical : « Thus, the word for 'sun' is grammatically feminine in German and Lithuanian, but masculine in Greek, Latin and the Romance languages, and neuter gender in Old Indic, Old Iranian and Russian. »

Notre recherche s'intéresse particulièrement, comme nous l'avons déjà dit, à l'assignation du genre grammatical du français. Holmes et Dejean de la Bâtie (1999, p. 479) décrivent le genre ainsi : « Gender refers to a system in which every noun falls into a different class, with class membership determining the form of other words that modify it or substitute it. » Cette définition convient très bien au français, comme nous allons le voir.

1.4 Le français et le genre grammatical

« In French, there is no neuter noun gender, so gender in French is the classification of nouns as masculine or feminine » (Surridge, 1995, p. 1). Le neutre aurait historiquement fusionné principalement avec le masculin, à partir du latin vulgaire (Hellinger et Bussmann, 2003, p. 90). De plus, le genre masculin est non marqué, selon Dubois et al. (2001, p. 294). C'est donc de ces deux genres dont il sera question tout au long de notre étude.

1.4.1 Quelques distinctions

Nous allons ici définir quelques termes afin de permettre au lecteur une bonne compréhension des détails de notre recherche.

1.4.1.1 L'assignation et l'accord

En français, le genre est une propriété du nom, c'est-à-dire que c'est celui-ci qui possède un genre. Par contre, c'est par le biais de mots d'autres catégories grammaticales qu'on a des indices de genre pour un nom donné. Ces mots d'autres catégories sont principalement les déterminants, les adjectifs, les participes passés ou présents agissant comme adjectifs et certains pronoms. « On a formal level, agreement establishes a syntactic relationship between a noun's satellite element, e.g. an article, adjective, pronominal or verbal form, and the noun's gender class » Hellinger et Bussmann (2001, p. 13). De plus, pour ce qui est des déterminants et des pronoms personnels, les formes plurielles ne marquent pas le genre. Déjà, Hockett (1958, p. 231) définissait le genre sensiblement de la même façon que Hellinger et Bussmann, comme des classes du nom reflétées dans le comportement de mots associés à celui-ci.

Il est donc important de faire une distinction entre assignation du genre à un nom et marques d'accord du genre (ou accord tout court). Pour le nom commun *table*, l'assignation du genre serait de savoir s'il est masculin ou féminin. Dans *la petite table*, *la* et *petite* ont des marques d'accord qui permettent d'identifier le genre du nom *table*, auquel on assigne le féminin (voir, entre autres, Sabourin, Stowe et de Haan, 2006). L'assignation du genre correspond donc au genre que l'on attribue à un nom, indépendamment des mots

périphériques du contexte phrastique de celui-ci. Notre recherche se concentre sur l'assignation du genre, et non sur l'accord en genre.

1.4.1.2 Le genre naturel et les épïcènes

Bien que pour certains éléments du lexique le genre soit un genre dit « naturel » parce qu'il est associé au sexe (un homme, une femme), il n'a habituellement aucun lien avec le sexe (Ibrahim, 1973, p. 12). De plus, certains noms sont dits « épïcènes » lorsqu'ils sont animés. Pour les noms animés humains, *épïcène* signifie qu'ils ont la possibilité qu'on leur assigne le genre masculin ou féminin, vu que la même forme correspond aux deux sexes (un/une enfant). Par contre, pour les noms animés animaux, *épïcène* signifie que les deux sexes sont associés à la même forme et au même genre grammatical (la souris). D'autres caractéristiques requièrent notre attention. Certains noms d'animaux ont une forme spécifique selon le sexe (le coq / la poule). Cependant, on peut aussi retrouver des noms humains à genre unique et qui représentent les deux sexes (une personne).

De façon générale, c'est tout de même le genre des noms animés qui a habituellement un lien avec le sexe. Pour ce qui est du genre des noms inanimés, l'assignation du genre semble sémantiquement arbitraire.

1.4.1.3 Les noms à initiale vocalique (NIV) et le français populaire du Québec

Le français populaire du Québec est une variante dialectale du français qui, bien sûr, possède ses caractéristiques propres. Comme la présente recherche a été effectuée à Montréal (Québec), il est important de parler de certains aspects qui ont trait à cette variante dialectale.

D'après Barbaud, Ducharme et Valois (1982, p. 104), on a tendance à féminiser les noms à initiale vocalique (NIV) en français populaire du Québec (*un avion* devient souvent *une avion*, par exemple). Cela est vérifiable statistiquement et par les règles d'accord (Barbaud, Ducharme et Valois, 1982, p. 105). Les noms féminins qui sont masculinisés sont beaucoup moins nombreux et sont surtout dus à la neutralisation phonologique du déterminant *un, une* en *enn/enne* (*enn homme / enne omelette*). Ces deux tendances à la féminisation et à la masculinisation « ne sont pas linguistiquement significatives au même degré [...] » (Barbaud, Ducharme et Valois, 1982, p. 104).

De plus, comme le genre d'un NIV n'est pas distingué dans plusieurs contextes d'accord (*l'avion, mon amie*), et ce, même au singulier, l'input des marques d'accord est moins fréquent pour ces noms (Holmes et Dejean de la Bâtie, 1999, p. 481). Il convient donc de ne pas traiter les NIV de la même façon que les non-NIV lorsqu'on veut étudier le genre, son assignation et son accord.

1.4.1.4 Autres éléments prédictifs pour l'assignation du genre des noms

D'autres caractéristiques influencent le genre des noms. Par exemple, les noms de certaines classes sémantiques possèdent toujours le même genre (SurrIDGE, 1995). Les jours de la semaine sont effectivement toujours masculins (le dimanche), tandis que les noms de disciplines académiques sont habituellement (mais pas toujours) féminins (l'astronomie, la chimie, etc.), notamment. Dans ces classes sémantiques à genre unique (ou presque unique), le genre n'a plus nécessairement de lien avec la terminaison des noms (bien que ce soit souvent possible), ce dont nous allons discuter dans la prochaine sous-section.

De façon générale, il semblerait aussi que la fréquence d'occurrence des noms influence l'assignation ou l'accord en genre. Selon Holmes et Dejean de la Bâtie (1999, p. 480): « [...] knowledge of the gender of a noun would gradually arise as a function of the speaker's frequent encounters with the noun in its various gender-specifying contexts. » Il est donc impératif de nous assurer que cette variable soit contrôlée ou, du moins, prise en compte lors de notre expérimentation.

1.4.2 Les terminaisons des noms

Il y aurait cependant un dernier indice de genre des plus importants et des plus documentés. Cette dernière hypothèse est au cœur de nos préoccupations.

La croyance générale est que « French is often regarded as having one of the most opaque gender systems » (Bloomfield, 1933). Cette catégorie est souvent vue comme arbitraire. Cependant, selon plusieurs auteurs (Tucker, Lambert et Rigault, 1977; SurrIDGE, 1995; Holmes et Dejean de la Bâtie, 1999; Lyster, 2006), il semblerait qu'à l'intérieur même de la forme des noms du français il y aurait des indices de genre. En effet, on a constaté que la majeure partie des terminaisons des noms correspondaient à un genre spécifique. Il y a eu

longtemps débat à savoir si c'était les suffixes dérivationnels (*-ation*, *-aison*, *-isme*, etc.) ou les terminaisons (pas nécessairement découpées aux suffixes dérivationnels) qui correspondaient le mieux au genre des noms, mais on en a conclu que les terminaisons englobaient les suffixes (du point de vue de leur efficacité à prédire le genre), et pouvaient donc décrire un plus vaste ensemble en ce qui concerne le genre. De plus, Mel'čuk (1974) affirme que « [...] the rule (based on 6,000 nouns) according to which words ending in a pronounced consonant are feminine and those ending in a pronounced vowel are masculine, applies in 85 % of cases. » Cependant, il semblerait que les terminaisons écrites prédisent le genre de façon plus exhaustive que les terminaisons orales. Holmes et Dejean de la Bâtie (1999, p. 482) proposent que « Spoken forms are more ambiguous than written forms : many spoken endings of French words can be spelled in different ways, and frequently each spelling is more typical of one gender than the other. »

Les terminaisons des noms n'éliminent pas complètement le problème de l'assignation du genre des noms en français. La variété des terminaisons est substantielle; le nombre des exceptions apparentes l'est aussi. De plus, il est difficile de savoir à partir d'où, dans le nom, une terminaison doit être prise en compte. Par exemple, dans le mot *terminaison*, est-ce qu'on doit considérer que c'est un nom se terminant par *-aison*, *-son* ou *-on* ? Cependant, cette possibilité d'un lien entre les terminaisons des noms et le genre de ceux-ci permet de cesser de croire que le genre grammatical est une donnée arbitraire du nom en français. Si le genre grammatical des noms inanimés semble vraiment sémantiquement arbitraire, il ne l'est probablement pas formellement. Généralement, c'est la forme des noms qui détermine leur genre, et non leur contenu (leur sens). « Thus, masculine and feminine are expressed either phonologically and/or morphologically on the noun itself [...] » (Hellinger et Bussmann, 2003, p. 91).

1.5 Le transfert

Vu que, pour notre recherche, nous serons en présence de deux langues fort différentes par rapport au français : l'une (l'espagnol) qui constitue une langue très parente du point de vue du genre et du lexique et l'autre (le chinois) qui n'a pas de genre du tout et qui n'a à peu près aucun congénère (mots similaires) avec le français, la comparaison de l'assignation du genre à des mots français de la part d'apprenants adultes du français langue seconde,

locuteurs soit de l'espagnol soit du chinois nous permet de vérifier le bien-fondé d'une stratégie particulière d'assignation, c'est-à-dire le recours au transfert. Avec la stratégie du recours aux terminaisons, nous rappelons que ce sont les deux principales hypothèses que nous voulons vérifier dans ce mémoire.

Nous allons d'abord brosser un tableau général de la notion de transfert. Ensuite, nous verrons ce qui a été écrit sur ce sujet, en relation au genre grammatical. Finalement, nous verrons en détail la notion de « congénère », ses définitions et son rôle dans le transfert. Cette notion aura une grande importance dans notre recherche.

1.5.1 Le transfert en général

D'après les observations de multiples chercheurs en acquisition et en didactique des langues, ainsi que de celles des enseignants du français langue seconde, les erreurs des apprenants ne sont pas toujours les mêmes d'un apprenant à l'autre, mais surtout d'un groupe linguistique à l'autre. Certains aspects de l'acquisition d'une langue semblent être constants à travers les divers groupes linguistiques, ce qui appuie l'idée d'une progression naturelle et d'un ordre naturel d'acquisition des éléments linguistiques au fur et à mesure que l'interlangue se développe. Cependant, d'autres aspects linguistiques semblent s'acquérir de manière beaucoup plus hétérogène d'un groupe linguistique à l'autre. Ces derniers aspects paraissent dépendre beaucoup plus d'un transfert de la L1 vers la L2 que d'une progression universelle de l'interlangue. Nous définissons le transfert, de manière générale, comme le recours à la langue première d'un locuteur donné afin de comprendre ou de produire un élément linguistique (ou parfois extralinguistique) d'une langue cible.

Odlin (1989, chap. 3) souligne quatre problèmes fondamentaux reliés à l'étude du transfert : le problème de la définition (qu'est-ce que le transfert?), celui de la comparaison (comment comparer les langues?, distance entre les langues), celui de la prédiction (comment interpréter des résultats supposément dus au recours au transfert?, transfert positif, négatif, etc.) et celui de la généralisation (comment généraliser l'interprétation de nos observations?). Nous nous attarderons d'abord et surtout au problème de définition.

Pour l'acquisition d'une langue seconde chez l'enfant, Dulay et Burt (1974, p. 52) disent que : « [...] universal cognitive mechanisms are the basis for the child's organization of the target language and [...] it is the L2 system [the target language] rather than the L1 system [the native language] that guides the acquisition process. » Globalement, cette affirmation est à l'opposé d'une théorie basée sur une prédominance de l'impact de la L1. Avant cette tendance à appuyer plutôt les mécanismes cognitifs universaux, il y avait prédominance de l'hypothèse du transfert dans le domaine de l'acquisition des langues secondes. Fries et Lado appuyaient avant tout le transfert. Déjà Fries (1949, p. 97) observait que « the problems of the Chinese student are very different from those of the Spanish speaker » en ce qui concerne leur acquisition de l'anglais. De plus, Lado (1957, p. 2) considère que : « [...] individuals tend to transfer the forms and meanings, and the distribution of forms and meanings of their native language and culture, to the foreign language and culture [...]. »

D'une étude à l'autre, le terme de transfert englobe différents éléments. Lorsqu'on parle de transfert, il est important de définir, d'une part, quelle est sa nature et, d'autre part, dans quelle mesure il a un rôle à jouer lors de l'acquisition d'une L2. Nous soutenons que le transfert est un processus qu'on ne peut pas observer directement. En effet, on ne peut que l'inférer par l'observation de ses résultats. De plus, lorsqu'on analyse ceux-ci, il est possible d'envisager un potentiel de transfert positif ou négatif. Un résultat issu d'un transfert positif est un bon résultat dû à une similitude de la L1 et de la L2 pour un élément linguistique donné. Un résultat issu d'un transfert négatif est inversement un mauvais résultat, donc une « erreur », dû à une différence de la L1 et de la L2 concernant un élément linguistique. Il nous semble cependant évident qu'il n'est possible d'avancer des hypothèses que sur le transfert négatif. Le supposé transfert positif pourrait être expliqué de multiples autres façons. En effet, qu'est-ce qui nous dit que la bonne assignation du genre pour un nom donné n'est pas simplement un signe que le locuteur de la L2 connaît tout simplement déjà le genre de ce nom, et ce, indépendamment de la similarité du genre dans sa L1?

Bien que le potentiel du transfert positif, à la base d'une non-erreur, soit possible, le transfert négatif, celui qui est à la base des erreurs, est plus décelable. Ellis (1994, p. 301) le confirme : « [I]n traditional accounts of language transfer, the research focus was placed on

the *errors* that learners produce. » Il semblerait que, dans l'ensemble des recherches sur les erreurs des apprenants, il y ait un continuum d'un impact plus grand de la L1 (dû au transfert) à un impact plus grand de la L2 (non dû au transfert). Bien sûr, il y a d'autres facteurs non linguistiques, tels que le degré de motivation, la personnalité de l'apprenant, le style cognitif, etc., qui influencent les processus d'acquisition, mais ceux-ci ne font pas l'objet de notre recherche. Cependant, tous les chercheurs s'entendent pour dire que le transfert n'est pas la seule raison qui puisse expliquer les erreurs commises par les apprenants : « Whereas Dulay and Burt (1973) report that transfer accounted for only 3 per cent of the errors in their corpus of Spanish-speaking learners' L2 English, Tran-Chi-Chau (1975) reports 51 per cent in adult, Chinese-speaking learners' English. » (Ellis, 1994, p. 302) La grande variation des résultats d'une recherche à l'autre, ou d'une L1 à l'autre, dépend du fait qu'il est très difficile de savoir si une erreur est due au processus du transfert (inter-linguistique) ou à un processus intra-linguistique.

Nous croyons que les deux hypothèses, celle des mécanismes cognitifs universaux et celle du transfert, pourraient avoir également un impact sur les erreurs ou les non-erreurs des apprenants, mais dans des domaines différents (bien que le flou de ces deux concepts ne nous permette pas de l'affirmer fermement). Par exemple, le niveau de langage ou l'élément linguistique pris en compte fait une différence. Ellis (1994, p. 316) parle de prévalence du transfert de la L1 notamment au niveau du lexique. Comme l'assignation du genre grammatical des noms est une question de lexique, nous émettons l'hypothèse que, pour cet aspect du langage, le transfert aurait un rôle particulièrement important à jouer.

Il y a eu beaucoup d'écrits scientifiques concernant le transfert puisque ce dernier fait partie des facteurs potentiels d'échec ou de réussite lors de l'acquisition d'une langue seconde. Comme on vient de le voir, les chercheurs ne s'entendent ni sur la définition du transfert ni sur son impact sur l'acquisition d'une L2. De manière générale, nous sommes d'accord avec Kellerman (1987, p. 3), qui définit le transfert comme suit : « [...] those processes that lead to the incorporation of elements from one language into another ». En effet, lorsque nous parlerons de transfert, nous nous entendrons pour une définition large, qui inclut tout aspect pris dans la L1 (ou langue de départ), puis reproduit dans la L2 (ou langue cible). De plus, Odlin (1989, p. 27) définit le transfert comme suit : « [T]ransfer is the

influence resulting from similarities and differences between the target language and any other language that has been previously (and perhaps imperfectly) acquired. » Cette définition donne un peu plus de détails par rapport à celle de Kellerman. Elle implique que les similarités (transfert positif) comme les différences (transfert négatif) peuvent susciter le transfert. Elle apporte aussi l'idée que le transfert peut provenir d'une ou de plusieurs langues apprises, plus ou moins bien, au préalable².

1.5.1.1 Le transfert superficiel et le transfert profond

Sabourin, Stowe et de Haan (2006), dans leur expérience sur la connaissance du genre grammatical en néerlandais par différents groupes linguistiques, décrivent deux types de transfert : le transfert superficiel et le transfert profond. Pour le genre grammatical, le transfert superficiel implique que la distribution du genre d'une langue à l'autre est similaire et qu'il y a de fréquents noms qui sont congénères. Le transfert profond est plus globalement le fait que le genre grammatical existe dans une langue donnée et que cet aspect linguistique est partagé avec une autre langue (par exemple, le français et le russe possèdent tous les deux le genre, même si en français le genre est limité au masculin et au féminin alors qu'en russe, il y a également un genre neutre). Le contexte linguistique du transfert superficiel implique donc celui du transfert profond, mais pas l'inverse.

1.5.2 Le transfert et le genre grammatical

« Many language teachers and linguists have believed that similarities and dissimilarities in word forms, along with similarities and dissimilarities in word meanings, play a major role in how quickly a particular foreign language may be learned by speakers of another language [...] » (Odlin, 1989, p. 77). Nous croyons que ces ressemblances ou dissemblances concernant le lexique peuvent influencer fortement, non seulement l'acquisition du vocabulaire, mais aussi celle du genre grammatical, élément grammatico-lexical. Adjémian (1983, p. 254), décrit le phénomène ainsi :

² Dans les données sociobiographiques récoltées auprès des participants de notre recherche, nous avons tenu compte des autres langues apprises antérieurement, mais seulement à titre de référence en cas de résultats marginaux. Nous croyons que, de toute façon, l'influence majeure, en ce qui concerne le transfert, est la L1.

[I]n the task of acquiring lexical knowledge in the target language, the learner may compare the lexical relatedness possibilities expressed in the L1 lexicon with the incoming target language lexical data, exploiting any that appear to match. This means that wherever it appears to fit, the learner may transfer a lexical rule or lexical feature from L1 into the learner-grammar lexicon. If this is the case, it should be possible to find in the learner's performance, or better still, in the learner's intuitions in the TL, evidence of L1 lexical properties.

Il est étonnant de constater que peu de recherches ont été effectuées sur le rôle du transfert dans l'assignation du genre. À part la recherche de Sabourin, Stowe et de Haan (2006), la plupart des articles que nous avons consultés parlent de l'effet des terminaisons sur l'assignation du genre ou parlent de l'accord en genre.

1.5.3 Les congénères et le transfert

C'est le transfert superficiel, tel que décrit par Sabourin, Stowe et de Haan (2006), qui nous intéresse surtout. Ce potentiel de recours à la stratégie du transfert demande que les deux langues en jeu possèdent non seulement les mêmes genres, mais aussi de grandes similarités lexicales, du point de vue de la forme. Ici, l'espagnol et le français sont donc de bons candidats pour notre recherche. En effet, s'il y a beaucoup de noms congénères d'une langue à l'autre, peut-être cela influencera-t-il les locuteurs d'une langue à avoir recours au transfert de leur L1 pour assigner le genre grammatical dans la L2. La distinction entre transfert superficiel et transfert profond ne nous aura seulement permis de clarifier notre cas particulier. Cependant, la comparaison des deux sortes de transfert ne sera pas l'objet de notre recherche.

Définissons maintenant la notion de *congénère*.

1.5.3.1 La définition de « congénère »

Lorsqu'un mot d'une langue donnée possède une grande similitude de forme avec un mot d'une autre langue, on dit qu'ils sont congénères interlinguaux, « la ressemblance formelle étant la condition première de la congénérarité [...] » (Theophanous, 2000, p. 95). *Congénère* veut donc dire ressemblant du point de vue de la forme, mais il y a tout un continuum de ce qui est complètement différent à ce qui est identique. Il est donc parfois difficile de déterminer si un mot est congénère ou non avec un autre mot.

En ce qui concerne des critères de sens, Granger (1993) distingue les bons congénères (good cognates) des mauvais congénères (deceptive cognates). Les bons congénères sont ce qu'on pourrait appeler les « vrais amis » (sompoteux (fr.) = *somptuous* (ang.)), c'est-à-dire qu'ils possèdent autant une forme qu'un sens très similaires, tandis que les mauvais congénères sont ce qu'on appelle traditionnellement les « faux amis » (actuel (fr.) ≠ *actual* (ang.)), c'est-à-dire les noms qui se ressemblent au niveau de la forme, mais qui diffèrent de sens. De plus, Granger sépare les mauvais congénères en deux sous-catégories : les mauvais congénères totaux et partiels. Les mauvais congénères partiels sont des mots qui sont de faux amis pour certaines acceptions seulement (expérience = *experience*, mais parfois *experiment*); les mauvais congénères totaux sont de faux amis pour toutes les acceptions (actuellement ≠ *actually*) .

Comment maintenant classifier les congénères, qu'ils soient bons ou mauvais, selon différents critères de forme? Lorsqu'on observe les données, on voit bien qu'il y a des congénères identiques, d'autres très similaires et d'autres de plus en plus éloignés en ce qui concerne leur forme, jusqu'à la perte de leur reconnaissance. Bien sûr, plus on a de connaissances de la langue cible, de la phonologie ou de l'étymologie, plus il est possible de reconnaître des similarités parfois obscures. Il y a des mots qui sont graphiquement identiques ou similaires, mais qui se prononcent différemment (don (français) / *don* (espagnol)), mais aussi des mots qui s'écrivent et se prononcent de façon identique (miel (français) / *miel* (espagnol)). Traditionnellement, lorsqu'on parle de congénères, on parle de mots écrits (homographes ou non). Il y a donc une multitude de possibilités quant au classement ou au choix des congénères en relation à leur forme (nombre de lettres ou de phonèmes communs; début, milieu ou fin de mot de la différence observée, distance phonologique des sons différents; longueurs des mots; etc.) (Theophanous, 2000, p. 96), mais là n'est pas l'objectif principal de notre recherche.

1.5.3.2 L'impact des congénères sur le transfert

Il va sans dire que la congénarité entre deux mots interlinguaux est sujette à influencer des résultats d'assignation de genre. Le fait de présenter à des locuteurs d'une langue donnée de vrais ou de faux amis (d'une L2 par rapport à la L1 des locuteurs) est susceptible d'influencer leurs décisions d'assignation du genre. De plus, même pour les vrais

amis, il est vraisemblable de voir un piège dans les congénères qui ne possèdent pas le même genre. Nous élaborerons ce point de vue dans la section 1.8 (Les questions de recherche et les hypothèses).

Les congénères peuvent fournir non seulement des informations sémantiques, mais aussi des informations morphologiques et syntaxiques. De plus, pendant que certaines informations peuvent induire en erreur, d'autres peuvent faciliter l'acquisition (Odlin, 1989, p. 83). Ard et Homburg (1983, p. 157) proposent que de plus grandes similarités entre deux langues que celles entre deux autres langues mènent toujours à des différences de résultats significatives, en ce qui concerne les résultats à des tests standards. Finalement, selon Ellis (1994, p.304), l'effet facilitateur de la L1 est évident dans différents aspects de l'acquisition d'une L2. Dans plusieurs cas, il est évident que, lorsque les langues concernées partagent un bon nombre de congénères (par exemple, l'anglais et le français), l'apprenant a une longueur d'avance pour l'apprentissage du vocabulaire. Ces trois références appuient notre hypothèse que la congénarité des noms pourrait largement influencer l'assignation du genre grammatical. Cependant, comme le dit Odlin (1989, p. 79), de plus en plus de recherches en sémantique lexicale contrastive montrent que la reconnaissance des congénères est le plus souvent un problème. Les apprenants ne notent pas toujours les similarités lexicales et ne croient pas toujours qu'il y a une relation de congénarité entre les mots comparés. Nous aurons l'occasion, dans notre recherche, de vérifier cette affirmation. Nous croyons tout de même que la reconnaissance des congénères est fonction de la similitude de forme. Certains congénères possèdent une similarité de forme plus éloignée, ce qui les rend plus difficiles à repérer. Comme le mentionnent Ard et Homburg (1983, p. 171), les résultats dépendent d'une matrice de similarités entre les items lexicaux de la L1 et de la L2. Dans notre étude, nous n'allons utiliser que des congénères très transparents du point de vue de la forme, et que de vrais amis.

1.6 Deux études importantes

Nous décrirons maintenant deux études sur lesquelles nous appuierons la majeure partie de notre recherche. Tout d'abord, l'étude de Sabourin, Stowe et De Haan (2006) sera notre point d'appui principal pour ce qui est de l'effet du transfert sur l'assignation du genre.

Ensuite, l'étude de Lyster (2006) sera notre point de départ en ce qui concerne l'effet des terminaisons des noms sur l'assignation du genre.

1.6.1 Sabourin, Stowe et De Haan (2006)

Sabourin, Stowe et de Haan (2006) comparent les résultats d'assignation du genre et d'accord en genre pour trois groupes d'apprenants du néerlandais : des locuteurs de l'allemand (25 participants), des locuteurs de trois langues romanes : le français, l'espagnol et l'italien (21 participants) et des locuteurs de l'anglais (24 participants). Cela donne donc un total de 70 participants. Nous ne parlerons ici que de la section sur l'assignation du genre, celle qui nous concerne.

Le néerlandais possède deux genres : le commun et le neutre. Les sujets parlant allemand avaient un potentiel de transfert superficiel et profond puisque leur langue possède beaucoup de congénères avec le néerlandais ainsi que trois genres, le masculin, le féminin et le neutre, qui correspondent à peu près à la distribution du genre en néerlandais (masculin et féminin = commun, neutre = neutre). Pour ce qui est des sujets parlant les langues romanes, seul le transfert profond était possible, mais pas le transfert superficiel puisque ces langues possèdent deux genres, le masculin et le féminin, qui ne correspondent pas à la distribution du néerlandais (masculin et féminin \neq commun et neutre). Finalement, l'anglais ne possède pas de genre grammatical, donc ne permettait aucunement à ses locuteurs d'avoir recours au transfert, que ce soit superficiel ou profond, lorsqu'ils ont pour tâche d'attribuer un genre à des substantifs néerlandais.

La recherche consistait à vérifier les résultats d'assignation du genre selon, tout d'abord, ces trois groupes linguistiques d'apprenants. Ensuite, les résultats ont été corrélés à deux niveaux de fréquence d'occurrence des noms (haute fréquence et moyenne fréquence), afin de vérifier si la familiarité d'un nom pouvait jouer sur l'assignation du genre. Finalement, l'assignation se faisait à l'aide des déterminants *de* et *het*, le premier (*de*) correspondant au genre non marqué, le commun; et le second (*het*) correspondant au genre marqué (de moins grande fréquence), le neutre. Tous les sujets étaient considérés de niveau avancé en néerlandais, et ce niveau a été testé au préalable (au-dessus de 95 % dans les tests de compétence en néerlandais). Les noms ont été choisis de façon à ce que les sujets les

connaissent, autant que possible. Cent soixante noms ont été testés dans l'expérience. Lorsque, par hasard, un sujet ne connaissait pas le sens d'un nom, on lui demandait d'encercler celui-ci. Par la suite, tous les noms encerclés ont été codés « non répondus ». Voici les résultats obtenus :

Tableau 1.3
Moyennes des résultats (et écarts) pour chaque groupe langagier, par type d'item
(tiré et traduit de Sabourin, Stowe et de Haan, 2006, p. 10)

	Haute fréquence		Fréquence moyenne	
	items en <i>de</i>	items en <i>het</i>	items en <i>de</i>	items en <i>het</i>
	Moyennes			
Allemand	99,3 % (95–100 %)	96,0 % (87–100 %)	96,0 % (90–100 %)	90,1 % (76–100 %)
Langues romanes	94,8 % (90–100 %)	92,8 % (72–100 %)	93,0 % (50–100 %)	76,5 % (52–98 %)
Anglais	86,8 % (65–100 %)	86,5 % (62–98 %)	88,6 % (70–98 %)	69,8 % (45–93 %)

Les résultats suggèrent que la fréquence d'occurrence des noms a un impact sur l'assignation du genre et qu'il y a une tendance à assigner le genre non marqué par défaut. En effet, les réponses des trois groupes linguistiques sont supérieures lorsque ces derniers sont soumis à des noms de haute fréquence d'occurrence, tandis que le déterminant correspondant au genre commun (« de »), le genre non marqué, est assigné par défaut lorsque les noms présentés sont de moyenne fréquence d'occurrence (voir le Tableau 1.3 et la Figure 1.1). De manière générale, les résultats des germanophones sont supérieurs à ceux des locuteurs de langues romanes, et les résultats de ces derniers sont supérieurs à ceux des anglophones. Cependant, la distance entre les résultats des germanophones et des locuteurs de langues romanes est beaucoup moins importante que la distance entre les résultats de ces derniers et ceux des anglophones. Cela suggère que le recours au processus du transfert a un impact significatif sur des résultats d'assignation du genre en néerlandais, mais que la distinction entre transfert superficiel et transfert profond est de moindre importance.

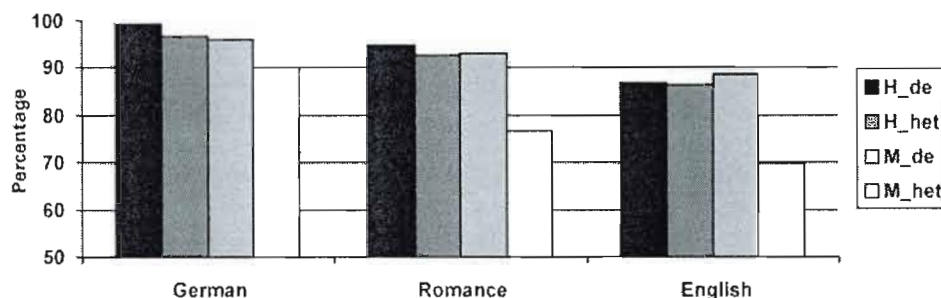


Figure 1.1 Résultats pour l'expérience I
(tiré de Sabourin, Stowe et de Haan, 2006, p. 11)
H = haute fréquence, M = moyenne fréquence

L'étude de Sabourin, Stowe et de Haan (2006) est donc un bon point de départ pour ce qui est de nos questions de recherches et nos hypothèses sur le transfert et l'assignation du genre grammatical.

1.6.2 Lyster (2006)

L'étude de Lyster (2006), quant à elle, se voulait une version plus moderne d'un aspect de la recherche de Tucker, Lambert et Rigault (1977). Dans cette dernière, tous les noms du Petit Larousse avaient été classés selon leur terminaison orale et écrite, et selon leur genre grammatical. Ensuite, plusieurs études avaient été menées sur de jeunes locuteurs natifs du français, sélectionnés dans diverses écoles. À l'aide de ces études, les auteurs voulaient vérifier le taux de bonnes assignations du genre par des locuteurs natifs à partir, notamment, de mots de très basse fréquence d'occurrence en français ou des non-mots à terminaisons françaises.

Lyster (2006) a refait le classement lexicologique des noms selon leur terminaison orale et écrite, puis selon leur genre, de Tucker, Lambert et Rigault (1977), mais sans les études subséquentes, faites sur des locuteurs natifs. De plus, il a basé son classement à partir du Robert Junior Illustré (9 961 noms, éliminant certains noms où le genre n'était pas relié à leur terminaison (par exemple, classes sémantiques à genre unique, comme décrit plus haut)), un corpus qui, selon l'auteur, était plus représentatif du vocabulaire potentiel d'un apprenant du FLS. De plus, dans sa recherche, Lyster (2006, p. 73) dit des terminaisons des noms

qu'elles se réfèrent à « variable orthographic representations of rhymes, defined as either a vowel sound (i.e., a nucleus) in the case of vocalic endings, or a vowel-plus-consonant blend (i.e., a nucleus and a coda) in the case of consonantal endings. »

Les résultats ont révélé que 81 % de tous les noms féminins et 80 % de tous les noms masculins du corpus possédaient une terminaison qui prédisait fortement le genre, ce qui contredisait la tradition grammaticale (de la classe d'enseignement) et qui suggérait des implications pédagogiques positives.

À l'aide du corpus, Lyster (2006, p. 74) a observé (comme bien d'autres auteurs) qu'une tendance se dessinait et permettait d'affirmer qu'une grande partie des phonèmes vocaliques finaux correspondaient à des noms masculins (phonèmes correspondant aux terminaisons écrites *-a*, *-eau*, *-oi*, *-un*, etc.). Inversement, une grande partie des phonèmes consonantiques finaux correspondaient à des noms féminins (phonèmes correspondant aux terminaisons écrites *-ange*, *-enne*, *-ette*, *-onte*, etc.). Cependant, l'auteur affirme, à l'instar de Holmes et Dejean de la Bâtie (1999, p. 482), que : « [...] final phonemes on their own are not consistently the most reliable predictors; instead, they interact with orthographic representations, which surpass phonological representations in terms of predictive value » (Lyster, 2006, p. 74). De plus, bien que plusieurs terminaisons correspondent à des suffixes dérivationnels, « [...] orthographic representations of rhymes proved instead to be considerably more productive than derivational suffixes in their capacity to account for a greater number of nouns and to generate a comprehensive set of rules governing gender attribution » (Lyster, 2006, p. 85).

C'est donc à partir de ces observations que Lyster (2006) classe les terminaisons écrites des noms du corpus selon trois groupes de terminaisons : typiquement féminines, typiquement masculines et typiquement ambiguës. Les terminaisons, pour chacun des groupes, devaient représenter 90 % et plus du corpus pour être classées comme typiquement féminines ou masculines. Celles qui n'obtenaient pas ce résultat faisaient partie des terminaisons typiquement ambiguës.

Tableau 1.4

Terminaisons typiquement féminines (tiré de Lyster, 2006, p. 87)

Vocalic

- -aie, -oue, -eue, -ion, -té, -ée, -ie, -ue

Consonantal

- -asse, -ace, -esse, -ère, -aïsse, -isse/-ice, -ousse, -ance, -anse, -eue, -once
- -enne, -onne, -une, -ine, -aine, -eine, -erne
- -ande, -ende, -onde, -ade, -ude, -arde, -orde
- -euse, -ouse, -ase, -aise, -èse, -oïse, -ïse, -yse, -ose, -use
- -ache, -iche, -èche, -oche, -uche, -ouche, -anche
- -ave, -ève, -ive
- -ière, -ure, -eure
- -ette, -ête, -ète, -atte, -otte, -oute, -orte, -ante, -ente, -inte, -ome
- -alle, -elle, -ille, -olle
- -aille, -eille, -ouille
- -appe, -ampe, -ombe
- -igue

Tableau 1.5

Terminaisons typiquement masculines (tiré de Lyster, 2006, p. 87)

Vocalic

- -an, -and, -ant, -ent, -in, -int, -om, -ond, -ont, -on (but not after s/ç)
- -eau, -au, -aud, -aut, -o, -os, -ot
- -ai, -ais, -ait, -ès, -et
- -ou, -oût, -out, -oux
- -i, -il, -it, -is, -y
- -at, -as, -ois, -oit
- -u, -us, -ut, -eu
- -er, -é after C (C≠t)

Consonantal

- -age, -ège, -ème, -ome/-ôme, -aume, -isme
- -as, -is, -os, -us, -ex
- -it, -est
- -al, -el, -il, -ol, -eul, -all
- -if, -ef
- -ac, -ic, -oc, -uc
- -am, -um, -en
- -air, -er, -erç, -ert, -ar, -arc, -ars, -art, -our, -ours, -or, -ord, -ors, -ort, -ir, -oir, -eur (if animate)
- -ail, -eil, -euil, -ueil
- -ing

Finalement, Lyster (2006) conclut avec des suggestions pédagogiques. Parmi un total de 400 terminaisons répertoriées, celles qui sont représentées seulement par cinq noms ou moins dans le corpus entier sont éliminées, ce qui simplifie le corpus à des fins pédagogiques. Parmi les 239 terminaisons restantes (environ 79 typiquement féminines, 92 typiquement masculines et 68 typiquement ambiguës), il a été observé que les terminaisons finissant par un *-e* muet sont majoritairement féminines et que les terminaisons ne finissant pas par un *-e* muet sont presque exclusivement masculines (voir tableaux 1.4, 1.5 et 1.6).

Cependant, il y a beaucoup de terminaisons ambiguës qui possèdent une terminaison en *-e* muet. L'auteur nous dit (p. 88), finalement, concernant une potentielle application pédagogique : « [A] rule of thumb that proves more reliable, therefore, is that noun endings without a final mute *e* are strong predictors of masculine gender (notable exceptions being *-té* and *-ion*) ». De plus, Lyster avait déjà observé, dans deux études antérieures, que chez des enfants de 10-11 ans en immersion, les taux d'assignation du genre reflétaient une attribution abusive (par défaut) du genre non marqué, le masculin (Lyster, 2004a et 2004b). Cette dernière affirmation corrobore l'affirmation similaire de Sabourin, Stowe et De Haan (2006), concernant le néerlandais.

Tableau 1.6
Terminaisons typiquement ambiguës (tiré de Lyster, 2006, p. 88)

Vocalic
• <i>-a, -ia, -oi/-oix, -çon/-son</i>
Consonantal
• <i>-erce/-erse, -orce/-orse</i>
• <i>-aque, -èque, -ique, -oque</i>
• <i>-ane, -ène, -one, -oine</i>
• <i>-amine, -omme, -ame, -ime, -ume, -arme, -erme, -orme</i>
• <i>-uge, -ige, -erge, -ange, -onge</i>
• <i>-erbe, -ape, -ipe, -ope, -oupe, -aphe</i>
• <i>-ure, -orc, -oire, -ir, -ère, -re after V(C)C, -eur (if inanimate)</i>
• <i>-ède, -ide, -ode</i>
• <i>-ate, -ite, -ote, -ute, -ecte, -ulte, -este, -iste, -epte</i>
• <i>-ale, -èle/-êle, -ile, -ole, -(V)ule, -le after V(C)C (C≠l),</i>
• <i>-agne, -igne, -ogne, -aigne/-eigne</i>

L'étude de Lyster (2006) constitue donc un bon point d'appui en ce qui concerne nos questions de recherches et nos hypothèses (autres que celles qui sont reliées au transfert), que nous allons élaborer dans la section 1.8.

1.7 L'espagnol et le chinois

Dans cette recherche, comme nous allons évaluer les résultats à des tests d'assignation du genre par des locuteurs de l'espagnol et du chinois. Nous devons donc maintenant décrire un peu ces langues en ce qui a trait à notre intérêt, la classification nominale.

1.7.1 L'espagnol

« Students who already know one Romance language have an advantage when it comes to studying another, since the degree of cognate relationship between Spanish, Portuguese, Italian, Catalan, French, etc., is high. » Cette affirmation de Teschner (1987, p. 256) est certainement vraie pour ce qui est de la reconnaissance du vocabulaire. L'est-elle aussi du point de vue de l'assignation du genre?

L'espagnol est une langue romane, comme le français. Nous avons déjà mentionné plus haut que les deux langues possédaient les deux mêmes genres, le masculin et le féminin. Cependant, ces genres ne sont pas distribués de la même façon (*la planète* (fém.) / *el planeta* (masc.)). On retrouve aussi, en espagnol, l'accord en genre des autres catégories de mots (adjectifs, déterminants, etc.).

Selon Teschner (1987, p. 257), dans son étude comparative du genre des congénères entre l'espagnol et le français, 75,43 % des noms sont congénères. Parmi ceux-ci, 92,38 % ont le même genre. Seulement 7,62 % (sur les 746 noms listés dans Teschner) des congénères n'auraient pas le même genre qu'en français. Cette étude se base sur les noms inclus dans le *Larousse français-espagnol* de 1967, qui contenait près de 19 841 noms. Ces chiffres ne tiennent pas compte du ratio des noms non congénères ayant le même genre ou non. Il est important de mentionner que ces statistiques tiennent compte d'une définition très souple de la notion de congénère. Tous ne seraient pas d'accord pour dire que *accueil* (fr.) et *acogida* (esp.), par exemple, sont congénères (cet exemple fait partie de la liste de Teschner). Parmi

les congénères entre les deux langues, on retrouve de vrais amis, mais aussi de faux amis (voir la définition de ces termes dans la section 1.5.3.1).

On retrouve aussi d'autres similarités avec le français, qui ne seront pas particulièrement prises en compte dans cette recherche. Par exemple, l'espagnol possède aussi des NIV, des homophones, des homographes, etc. Souvent, comme pour le français, le genre grammatical de certains noms peut être expliqué par des règles sémantiques (ex. : les noms de rivières sont masculins), par des règles de forme grammaticale (ex. : les noms composés sont masculins), par des règles étymologiques (ex. : les noms d'origine grecque se terminant en *-a* sont masculins), par des règles phonologiques (ex. : les noms qui portent un *á* accentué sur la première syllabe sont masculins), etc. (Franceschina, 2005, p. 97). Finalement, comme en français, l'espagnol possède différentes règles de terminaisons nominales, qui parfois prédisent le genre grammatical. Ce sont probablement ces dernières règles qui sont celles qui prédisent le mieux l'assignation du genre grammatical en espagnol.

Harris (1991, p. 32) hiérarchise les classes de flexions nominales (terminaisons) de l'espagnol. En premier lieu, il y a les classes de noms à terminaisons régulières marquées pour le genre. En espagnol, on peut dire que, de façon générale, lorsqu'un nom se termine en *-o*, il est masculin (ex. : *grupo* (m)); lorsqu'il se termine en *-a*, il est féminin (ex. : *familia* (f)). En effet, selon Hellinger et Bussmann (2002, p. 254), les noms espagnols se terminant par *-o* sont à 99,89 % masculins, tandis que ceux se terminant par *-a* sont à 96,6 % féminins. Ensuite, il y a les classes de noms à terminaisons régulières non marquées pour le genre (ex. : *parte* (m/f), *par* (m/f)). Il faut noter que, si Harris considère la terminaison *-e* comme non marquée pour le genre, Hellinger et Bussmann disent que les noms qui ont cette terminaison sont tout de même à 89,35 % masculins. Finalement, il y a la classe de résidus irréguliers. Dans cette dernière classe, on retrouve, notamment, ce que nous appelons les noms à genre inversé. Dans le nom *dia* (jour), il y a une terminaison typiquement féminine, *-a*. Pourtant, ce nom est masculin. Dans le nom *mano* (main), il y a une terminaison typiquement masculine, *-o*. Pourtant, ce nom est féminin.

On peut aussi retrouver certaines terminaisons associées à celles du français. Par exemple, lorsqu'un nom français se termine en *-té*, il a de fortes chances que son équivalent

en espagnol se terminera en *-dad*. Cependant, Teschner (1987, p.257) nous révèle qu'il n'y aurait que peu de similarité entre les suffixes des congénères des deux langues. En effet, seulement 7,11 % des congénères auraient pratiquement le même suffixe dans les deux langues.

Comme pour le français, le genre grammatical des noms animés va de pair avec le sexe de ces noms (avocat : *abogado* / avocate : *abogada*). Cependant, un seul nom peut parfois représenter les deux sexes (le/la journaliste : *el/la periodista*) : ces noms sont épicènes. De plus, certains noms inanimés peuvent être associés à un genre dit naturel, vu les origines culturelle et historique. Selon Hellinger et Bussmann (2002, p. 254), dans *la muerte* (la mort), le concept est habituellement représenté par une femme, et ce, toujours comme en français.

Il semblerait que l'espagnol n'ait pas de suffixe dérivationnel masculin, mais qu'il en ait de féminins (*-isa, -esa, -iz, -ina, -sa*) (Hellinger et Bussmann, 2002, p. 255). Cela appuie donc, comme pour le français, que le masculin est la valeur non marquée, tandis que le féminin est la valeur marquée.

1.7.2 Le chinois

Pour sa part, le chinois est une langue à classificateurs sans genre grammatical (Hellinger et Bussmann, 2002, p. 27). Comme nous l'avons vu plus haut, cette langue utilise plutôt des classificateurs, ou compteurs, afin de classer le lexique nominal. De plus, il n'y a pas de noms congénères (ou presque pas) avec le français, du fait que cette langue n'est pas du tout de la même famille. Il y a donc peu de choses à dire sur cette langue quant aux aspects qui nous concernent. Cependant, il est intéressant de constater que, comme dans la plupart des langues du monde, on retrouve en chinois des traces, surtout à l'écrit, d'une attitude sociale en faveur du masculin (ou plutôt de l'homme) (Hellinger et Bussmann, 2002, p. 33). Cela pourrait influencer des résultats en faveur d'une assignation du genre grammatical masculin, mais cela serait certainement une influence minime.

1.8 Les questions de recherche et les hypothèses

Voici maintenant les principales questions de recherche que nous nous posons concernant l'assignation du genre en français par des locuteurs de L1 différentes (ici, des

hispanophones et des sinophones). Les questions de recherche se subdiviseront en trois groupes : transfert, terminaisons et autres. Ces questions seront, lorsque pertinent, suivies de nos hypothèses. Selon Franceschina (2005, p. 191), « adult L2 learners can acquire nativelike [sic] knowledge of grammatical gender » et « the learner's L1 plays a crucial role in determining whether nativelike [sic] attainment is possible in the area of grammatical gender. » Notre recherche nous permettra-t-elle de confirmer ces affirmations?

Notre question générale est la suivante :

Q1 : De façon globale, quel groupe linguistique aura plus de succès dans l'assignation correcte du genre en français, celui dont la L1 possède un genre, les hispanophones, ou celui dont la L1 n'en possède pas, les sinophones?

Ici, aucune hypothèse n'a été formulée. Nous pensons que les groupes ont des avantages différents. Les hispanophones ont l'avantage d'être en terrain connu (ils possèdent un genre), mais les sinophones ne peuvent pas être induits en erreur (ils n'ont pas de genre possiblement erroné à transférer).

1.8.1 Questions de recherche et transfert

En ce qui concerne l'influence du recours au transfert comme stratégie d'assignation du genre grammatical en français, voici les questions que nous nous posons :

Q2 : Dans quelle mesure y a-t-il recours au transfert par les hispanophones?

Q3 : Quel est l'impact, positif ou négatif, des congénères sur ce transfert?

Nous émettons l'hypothèse que les hispanophones auront grandement recours au transfert. Il nous sera difficile (mais pas impossible) de répondre à la question de l'influence du transfert positif. Il est trop difficile d'interpréter des résultats positifs; il vaut mieux observer les erreurs. En effet, si nous arrivions à de bons résultats chez les hispanophones pour l'assignation du genre aux noms et que ces résultats puissent être interprétés par la stratégie du recours au transfert, comment distinguer le recours à cette stratégie de la simple connaissance du genre de ce nom, sans recours à une stratégie? En fait, plusieurs

interprétations sont alors possibles. Il est donc impossible d'interpréter solidement ces données. Nous observerons donc uniquement les erreurs, impliquant que le participant a eu recours au transfert, mais que les résultats sont négatifs (donc une erreur).

Pour ce qui est de l'influence des congénères, nous croyons que cet aspect aura un grand impact, autant positif que négatif, pour les hispanophones. L'assignation du genre aux congénères qui ont le même genre dans les deux langues devrait être particulièrement bien réussie, tandis que l'assignation du genre aux congénères qui n'ont pas le même genre dans les deux langues devrait être particulièrement échouée.

Par exemple, si on présente le mot *miel* (français) à un hispanophone, nous émettons l'hypothèse que celui-ci assignera le genre féminin à ce mot. En effet, *miel* est féminin en espagnol, tandis qu'il est masculin en français. Une erreur de la sorte pourrait être plus facilement évitée dans le cas du mot *nettoyage*. Ce mot, qui est masculin en français, correspond au mot *limpieza*, féminin en espagnol. Bien que le mot correspondant en espagnol ne soit pas du même genre que le mot français, il nous semble fort probable que le taux de réussite d'assignation du genre (par les hispanophones) pour *nettoyage* sera plus élevé que pour le mot *miel*.

Pour ce qui est des non-congénères, nous pensons qu'il y aura un effet du recours au transfert, mais dans une proportion beaucoup moins grande. Pour les non-congénères qui auront le même genre dans les deux langues, l'assignation du genre devrait être mieux réussie, de façon générale, que pour les non-congénères qui n'ont pas le même genre dans les deux langues, mais avec une différence de résultats beaucoup moins grande que pour les congénères.

Les résultats d'assignation du genre par les sinophones ne seront influencés ni par la similarité de forme (congénarité) ni par la similarité de genre avec le français. Comme nous l'avons dit précédemment, la stratégie du recours au transfert de la L1 est impossible de la part du groupe des sinophones puisqu'ils ne connaissent pas le genre dans leur langue et n'ont pas de noms congénères avec le français. Donc, comme nous l'avons déjà mentionné, ce groupe servira de contrôle pour nos résultats hispanophones.

1.8.2 Questions de recherche et terminaisons

Lorsqu'on ne peut pas expliquer les résultats par le transfert, peut-être y a-t-il d'autres explications évidentes. Voici maintenant les questions concernant la stratégie du recours aux terminaisons :

Q4 : Est-ce que les terminaisons marquées pour le genre (typiquement féminines et typiquement masculines) suscitent de meilleurs résultats que les terminaisons non marquées pour le genre (typiquement ambiguës)?

Q5 : Est-ce que les noms qui ont une terminaison en *-e* muet final, à l'écrit, (indépendamment du fait que leurs terminaisons soient marquées ou non pour le genre) incitent les apprenants à assigner le genre féminin et est-ce que les noms qui n'ont pas de terminaison en *-e* muet final incitent les sujets à assigner le genre masculin?

Nos hypothèses concernant la stratégie du recours aux terminaisons comportent plusieurs volets et s'adressent autant aux hispanophones qu'aux sinophones. Les résultats des sinophones nous sembleront même plus appropriés puisque non influencés par un potentiel recours au transfert. Nous émettons l'hypothèse que la stratégie du recours aux terminaisons sera fortement présente, surtout chez les sinophones, qui n'auront pas pu faire appel au transfert. Les résultats pour les noms à terminaisons marquées pour le genre devraient être supérieurs puisqu'ils donnent un indice intrinsèque de genre. Cependant, la terminaison en *-e* muet (écrit) devrait avoir un plus gros impact sur l'attribution du genre. Lorsque les noms français se terminent par *-e*, on leur assignera le genre féminin. Lorsque les noms français ne se terminent pas par *-e*, on leur assignera le genre masculin. Cette stratégie devrait obtenir plus de succès que celle attribuant de meilleurs résultats aux noms à terminaisons marquées. En effet, il est beaucoup plus simple d'appliquer une stratégie binaire (*-e* / pas de *-e*) que d'apprendre le genre de chacune des terminaisons du français.

1.8.3 Autres questions de recherche

Pour compléter, d'autres facteurs, secondaires, mais non négligeables, doivent être pris en compte dans cette recherche. Voici les différentes questions qui y sont associées :

Q6 : Le taux d'assignation correcte du genre est-il fonction de la fréquence des noms présentés?

Q7 : Le taux d'assignation correcte du genre est-il fonction du niveau de connaissance du vocabulaire des apprenants?

Q8 : De façon générale, est-ce que les sujets ont tendance à assigner le masculin par défaut (genre non marqué)?

Q9 : Y a-t-il un lien entre le sexe des participants et leurs résultats (corrélation entre le sexe et l'assignation du genre, meilleurs résultats pour un sexe donné)?

Nous émettons l'hypothèse que, plus un nom est fréquent, plus l'assignation de son genre sera réussie, et ce, indépendamment des autres facteurs. Cela est appuyé par notre croyance que le genre est essentiellement appris au même moment que le nom; il est un aspect de l'unité lexicale. Cependant, il est possible de retrouver beaucoup de variabilité à cet effet puisque, selon la technique d'enseignement/apprentissage, on apprend le nom avec ou sans son genre. Nous croyons aussi que les apprenants avancés en connaissance du vocabulaire français auront plus de facilité à attribuer un genre aux noms puisqu'ils ont probablement eu plus longtemps d'influence de l'input.

Ensuite, nous émettons l'hypothèse que les erreurs des participants devraient tendre vers le masculin, ce genre étant le genre non marqué du français, selon Lyster (2006), notamment.

Pour ce qui est du lien avec le sexe des sujets, nous croyons qu'il pourrait y avoir un effet. Dans notre pratique de l'enseignement du français langue seconde à des anglophones, nous avons, de façon anecdotique, observé à deux reprises que les erreurs de genre correspondaient au sexe de l'apprenant. Dans ces cas, une femme mettait au féminin des noms masculins et un homme mettait au masculin des noms féminins. Bien que ce ne soit pas le cœur de nos préoccupations de nous pencher sur cette question, nous tenterons d'observer les résultats de notre recherche à titre exploratoire. Nous pensons que ces observations anecdotiques seraient susceptibles d'apparaître à nouveau comme stratégie potentielle.

Cependant, nous ne croyons pas que les hommes réussissent mieux l'assignation du genre que les femmes, et vice versa. La variable *sexe* ne pourra jouer que pour la distribution des erreurs, et non pour le taux de réussite global. Ces dernières affirmations sont nos hypothèses quant à la variable *sexe*.

1.9 Conclusion

Maintenant que nous avons terminé de broser le tableau du cadre théorique et posé nos questions de recherche, nous allons établir les bases méthodologiques de notre étude afin de tenter de répondre à ces questions de recherche et de confirmer ou d'infirmer nos hypothèses.

CHAPITRE II

MÉTHODOLOGIE

2.1 Introduction

Dans l'expérience que nous avons effectuée, nous n'avons testé que l'assignation du genre grammatical en français, et non l'accord en genre. De plus, nous nous sommes limité aux résultats d'assignation de genre à l'écrit seulement, et ce, chez deux groupes linguistiques différents.

Pour bien mener une recherche de ce type, il est important, donc, de bien choisir les participants de l'étude, mais aussi de bien choisir les mots qui seront testés. Une grande part de la recherche est de confectionner des tests qui permettront de répondre à nos interrogations, tout en éliminant toute variable inopportune. En effet, le problème de l'interprétation des résultats pourrait être un casse-tête des plus difficiles. Il est donc important, en plus d'utiliser un groupe contrôle, de bien déterminer les critères de sélection des mots. Pourquoi un mot plutôt qu'un autre? Les premiers critères doivent être les critères de similitude avec l'espagnol dont nous avons parlé plus haut : la question à savoir si le nom français choisi possède ou non le même genre qu'en espagnol et celle à savoir si ce nom est congénère ou non avec l'espagnol. En vertu de la notion du transfert, si un nom français particulier a le même genre que son équivalent sémantique en espagnol et qu'en plus sa forme est semblable à celle en espagnol (ils sont alors congénères), on devrait s'attendre à ce que les ressemblances entre l'espagnol et le français permettent à l'apprenant hispanophone du français de facilement et correctement assigner le genre correct au dit nom (ce qu'on appelle le « transfert positif »).

De plus, un contrôle particulier est nécessaire pour ce qui est des terminaisons des noms sélectionnés, et ce, afin d'éviter que ces terminaisons influencent les résultats

concernant le transfert. Finalement, la fréquence ou le degré de connaissance de ces noms par les apprenants du français doit être pris en compte pour ne pas obtenir des résultats inanalysables. Nous verrons les critères qui nous ont permis de choisir ou non un nom particulier. De plus, plusieurs caractéristiques nominales secondaires peuvent influencer le choix du genre des noms français lors de son assignation par des locuteurs non natifs. Il est donc important d'éliminer certains mots selon certains critères d'exclusion afin d'isoler le plus possible les variables qui nous intéressent.

2.2 Les participants

Nous avons recruté 60 participants qui ont pour langue maternelle l'espagnol ou le chinois (surtout le mandarin). Cet échantillon a été réparti assez équitablement chez les deux groupes linguistiques, à savoir 28 hispanophones et 32 sinophones. Nous avons principalement recruté les sujets à l'École de langues de l'UQAM; ils étaient donc tous des adultes et ils devaient connaître minimalement un peu de français. Ils ont été choisis parmi tous les niveaux de connaissance du français, de débutant à très avancé (donc, tous les groupes de l'École de langues étaient admissibles), mais étaient essentiellement intermédiaires ou avancés. Néanmoins, certains participants n'étaient pas de l'École de langue, mais étudiaient à l'UQAM. Pour la majorité des participants, nous les avons recrutés directement dans huit classes de français auxquelles nous avons eu accès grâce à quatre enseignants. Nous présentions brièvement (en cinq minutes) notre besoin de participants. Puis nous demandions aux intéressés de laisser leurs coordonnées sur une feuille. Quant aux participants n'étant pas des étudiants de l'École de langue, ils ont été abordés directement et individuellement. La répartition des compétences en français des participants était similaire entre les hispanophones et les sinophones.

Pour ce qui est du sexe des participants, nous avons obtenu (sans le chercher volontairement) un ratio de deux femmes pour un homme, ce qui correspond aux statistiques d'inscription de l'UQAM. Les 28 participants hispanophones étaient répartis en 19 femmes pour 9 hommes, tandis que les 32 sinophones étaient répartis en 22 femmes pour 10 hommes. Cela a donné un total, les deux groupes linguistiques confondus, de 41 femmes pour 19 hommes.

Nous n'avons pas tenu compte de la langue maternelle chinoise spécifique (mandarin, cantonais, szechuanais, etc.) dans notre sélection des participants. De la même façon, les hispanophones ont été acceptés, peu importe leur pays d'origine (Mexique, Espagne, Venezuela, etc.). Cependant, ces différentes données sociobiographiques ont été prises en note lors de la passation des tests afin de potentiellement nous aider lors de l'analyse des résultats.

Les hispanophones ont clairement un fort potentiel de recours au transfert, aussi bien superficiel que profond (voir p. 20, pour les définitions). En effet, l'espagnol possède les deux mêmes genres qu'en français, le masculin et le féminin. De plus, il y a beaucoup de congénères d'une langue à l'autre.

Les sinophones, quant à eux, n'ont aucun potentiel de recours au transfert. Le « chinois », terme qui inclut toutes les langues chinoises, ne possède aucun genre grammatical ressemblant au français, ni de congénères avec cette langue. Les sujets sinophones étaient donc, en grande partie, un groupe contrôle afin d'appuyer ou non l'hypothèse du recours au transfert des hispanophones, transfert donnant lieu à des résultats positifs ou négatifs. Par exemple, si certaines réponses de la part des hispanophones semblaient indiquer un potentiel de recours à une stratégie de transfert, cette hypothèse serait immédiatement et fortement affaiblie par des résultats similaires de la part des sinophones, puisque ces derniers résultats ne pourraient en aucun cas être dus à un recours au transfert. De plus, le groupe des sinophones nous a apporté la possibilité d'observer les résultats par rapport à l'influence des terminaisons, et ce, sans le parasitage d'un transfert linguistique possible. Nous avons donc choisi ces deux groupes linguistiques parce que les langues parlées par ceux-ci étaient, selon nous, les deux pôles d'un continuum de ressemblance/différence au français quant au genre grammatical.

Le classement des sujets pour l'expérience ne dépendait pas de leur niveau en français reconnu à l'École de langues ou du niveau auquel ils se croyaient être, ou qu'ils me disaient être. Ils ont été classés selon leur niveau de connaissance du vocabulaire, et ce, à partir d'un test de reconnaissance de mots. Ce test a permis de savoir si les sujets connaissaient le sens de noms présélectionnés dans le corpus de Jean Baudot (1992).

2.3 Le corpus de Baudot (1992)

Nous avons seulement choisi des noms qui se retrouvaient dans le corpus de Jean Baudot (1992) afin d'obtenir des items pour lesquels nous avons une fréquence d'occurrence documentée. Ce corpus de fréquence est constitué de 20 000 mots. Chacun des mots est classé selon sa catégorie grammaticale et selon un numéro représentant sa fréquence. Ces numéros sont répartis de un à environ 6 000 (la préposition *de*), et plus un mot a un numéro élevé, plus il est fréquent. Même si le niveau de reconnaissance des participants allait être évalué lors de nos tests, il nous fallait sélectionner des noms qui avaient des fréquences d'occurrence diversifiées afin de maximiser nos chances d'obtenir une répartition équitable entre des noms facilement et difficilement reconnus. De plus, cela nous a permis aussi de comparer la fréquence des noms du corpus aux taux de reconnaissance de ceux-ci lors du test de reconnaissance.

« Le corpus est composé de 803 échantillons de textes (d'environ 1 000 à 1 500 mots chacun) répartis en 15 genres littéraires, eux-mêmes divisés en sous-genres. Les dates de rédaction de ces textes s'échelonnent de 1906 à 1967, la majorité d'entre eux (726 sur 803) se situant entre 1960 et 1967 » (Baudot, 1992). Voici les origines nationales de ces textes en langue française : France (62 %), Canada (37 %), Autres pays (9 textes) (1 %). Pour ce qui est des types de publications, voici le partage : revues et magazines (42 %), livres et manuels (25 %), journaux (24 %), bulletins et rapports (7 %), brochures et circulaires (2 %). C'est un corpus représentatif du français écrit, ce qui nous convenait puisque les tests allaient être à l'écrit (affichage de mots à l'écran d'un ordinateur). De plus, il contient 21 684 types (mots), ce qui nous a donné un grand choix de mots.

2.4 La confection des tests

Deux tests ont été créés pour accomplir la recherche. Premièrement, les participants ont dû répondre à un test de reconnaissance de noms communs du français (test ON). Deuxièmement, ils devaient répondre au test principal, celui d'assignation du genre grammatical (test MF).

Nous décrivons ici la façon utilisée pour sélectionner les 75 mots du test principal, qui seront aussi utilisés, d’abord, pour le test de reconnaissance.

Premièrement, à l’aide du logiciel Microsoft Access, nous avons confectionné une base de données. Cette base de données nous a permis de classer chacun des noms communs selon une liste de critères précis afin de choisir les items du test principal. Les noms faisaient partie de classes hermétiques à critères croisés. Nous avons fait nos choix définitifs parmi la liste (base de données), qui a été préalablement épurée par des critères d’exclusion. Nous avons éliminé :

- 1) tous les NIV, vu que ces noms ont un input différent des autres (voir ci-haut, p. 12), et pour éviter toute confusion entre l’input « standard » et l’input québécois;
- 2) tous les noms commençants par la lettre « h » et dont l’« h » est aspiré puisqu’il est difficile, même pour les locuteurs natifs, de déterminer s’il est aspiré ou non (le homard, *l’homard), ce qui aurait rendu confuse la distinction entre les noms en « h » NIV et en « h » non-NIV, distinction importante pour notre recherche;
- 3) tous les noms animés, parce que les épicènes humains possèdent les deux genres, et pour éliminer la confusion potentielle avec le sexe (ou le genre naturel) (*un/une journaliste* et *un chien*, qui peut aussi faire référence tant au mâle qu’à la femelle);
- 4) tous les noms qui sont homographes avec d’autres noms en français et qui auraient un genre différent, pour éviter la possibilité d’assigner les deux genres (*le/la mémoire*);
- 5) tous les noms qui sont homophones avec d’autres noms français et qui auraient un genre différent, pour éviter l’influence possible de ce genre différent (*la mer* et *le maire*);

- 6) tous les noms composés, à cause de la difficulté de savoir à quel morphème le genre est associé ou à cause de morphèmes n'appartenant pas à la classe des noms, habituellement (un après-midi, un rendez-vous);
- 7) tous les noms qui ont un genre inversé en espagnol : *-o/-a* (Harris, 1991), parce qu'il aurait été difficile de déterminer si l'assignation du genre est due à un transfert du genre du nom ou du genre de sa forme inhabituelle (le problème = *el problema* (masculin) ou la main = *la mano* (féminin));
- 8) tous les faux amis du français pour un hispanophone (le sol = *el suelo*), mais pas l'inverse, les faux amis de l'espagnol pour un francophone (le soleil = *el sol*);
- 9) tous les noms qui font partie d'une classe sémantique à genre unique (SurrIDGE, 1995), parce qu'une potentielle connaissance de ces règles ne permet plus de suggérer qu'il y a transfert ou non et parce que le lien avec la terminaison est souvent annulé (le lundi, le rouge, la Ford);
- 10) tous les noms qui correspondent à des emprunts à terminaisons non françaises (le jazz), parce qu'il aurait été impossible de répondre aux questions de recherche qui touchent aux terminaisons;
- 11) tous les noms qui sont toujours au pluriel (les mœurs), parce que l'input de leurs marques d'accord est réduit;
- 12) toutes les abréviations lexicales (une photo, une radio), puisque leur genre est associable à la terminaison du mot non abrégé.

Une fois les noms éliminés selon les critères 1-12, nous avons ensuite établi le taux de fréquence de chaque mot selon Baudot (1992), afin de nous assurer que nous avions une distribution équitable entre mots fréquents et mots moins fréquents. Les noms ont été choisis de façon à créer un certain continuum des fréquences afin d'éviter de découper celles-ci en niveaux de fréquence élevée, moyennement élevée ou très élevée. Nous ne pouvions pas grouper les mots par taux de fréquence. Cela nous aurait donné, notamment, une variable

supplémentaire dont il nous aurait fallu tenir compte, ce qui n'aurait pas été possible statistiquement vu la faible quantité de mots choisis.

Par la suite, afin de nous assurer de l'exactitude des équivalents en espagnol, de leur genre, de l'existence de potentiels faux amis et de congénères, ou de mots similaires aux équivalents que nous avons trouvés (en consultant des dictionnaires bilingues), nous avons fait vérifier la liste des noms que nous avons choisis par deux locuteurs natifs de l'espagnol du département de linguistique et de didactique des langues (UQAM). Ce sont ces vérifications qui nous ont fait épurer notre liste de noms. Nous avons éliminé plusieurs mots préalablement choisis, pour continuer les recherches et enfin établir la liste finale. De plus, nous nous sommes aidé du Real Academia Española informatisé (RAEi), pour l'espagnol, et du Trésor de la Langue Française informatisé (TLFi), pour le français.

Tableau 2.1
Les 15 non-mots

Terminaisons		
Masculines (5)	Féminines (5)	Ambiguës (5)
blérisme	laitation	mâchoix
lipanage	caricité	fison
gontirail	cramance	bauxique
sartean	néméade	néramme
mitran	dérise	vérode

Mais, pourquoi 75 mots, précisément ? Tout d'abord, nous avons inclus 15 non-mots (voir Tableau 2.1), ce qui nous a permis de contrôler la validité des résultats. Les non-mots sont des mots que nous avons créés de toutes pièces à partir de nos intuitions des règles phonotactiques du français. De plus, nous nous sommes assuré qu'ils ne comptaient pas parmi les mots du TLFi (environ 800 000 mots).

Puisque nous étions intéressé à mesurer la connaissance du véritable lexique français des participants, nous avons pu contrôler celle-ci à l'aide des non-mots. Si, par exemple, certains sujets prétendaient connaître le sens de non-mots dans le test de reconnaissance, ils perdaient un point par non-mot. De plus, nous avons pu vérifier si le genre assigné aux non-mots correspond au genre de leur terminaison. Cela nous a donné des indices afin de répondre à certaines de nos questions de recherche concernant les terminaisons.

Nous nous sommes assuré que les radicaux et les terminaisons des non-mots correspondaient à la phonotactique du français et que les terminaisons étaient distribuées également selon les trois catégories de Lyster (2006) : typiquement masculine (ex. : blérisme = *-isme*), typiquement féminine (ex. : laitation = *-ion*) ou typiquement ambiguë (ex. : mâchoix = *-oix*). Cependant, nous n'avons pas pris compte de leur potentielle congénarité avec les noms de l'espagnol. De plus, elles devaient toutes être différentes bien que faisant toutes partie du français. Il y avait donc 60 noms réels plus 15 non-mots (cinq non-mots par catégorie de terminaisons), ce qui donnait 75 items.

Pourquoi 60 noms réels, alors? L'application des critères d'exclusion mentionnés plus haut nous a obligé d'éliminer un certain nombre de mots. Afin d'établir la liste finale de noms du test principal, nous avons par ailleurs utilisé les critères suivants (voir Tableau 2.3, pour la liste des mots) :

- 1) même genre ou non avec l'espagnol (classement binaire, mais donnant quatre regroupements selon la combinaison d'association des genres du français à l'espagnol = m/m - f/f - m/f - f/m; par exemple, le mot « miel » est masculin en français (le miel) et féminin en espagnol (*la miel*), il a donc été classé dans la catégorie m/f; le mot « nageoire » est féminin en français (la nageoire) et féminin en espagnol (*la aleta*), il a donc été classé dans la catégorie f/f, etc.;
- 2) congénère ou non avec l'espagnol (classement binaire = congénère - non congénère; dans notre même exemple, « miel » est congénère avec l'espagnol, tandis que « nageoire » ne l'est pas);

- 3) terminaison marquée ou non (classement binaire = marquée / non marquée; par exemple, le mot « miel » a une terminaison marquée pour le genre, *-el*, qui est typiquement masculine (selon Lyster, 2006); le mot « nageoire » a une terminaison non marquée pour le genre, *-oire*, qui est une terminaison ambiguë (toujours selon Lyster, 2006));
- 4) terminaison se terminant en *-e* (classement binaire = en *-e* / pas en *-e*; par exemple, le mot « miel » ne se termine pas en *-e*, mais le mot « nageoire », oui).

Les noms ont donc été méticuleusement répartis selon ces critères afin de pouvoir répondre à toutes nos questions de recherche et afin de nous assurer que le corpus représentait de manière adéquate le lexique du français.

Dans le Tableau 2.2, nous pouvons voir la répartition des terminaisons choisies. Celles-ci dépendaient plus des mots potentiels nous permettant d'obtenir au moins un exemplaire de chacune des catégories présentées par les quatre critères de sélection énoncés plus haut. Elles sont classées selon qu'elles sont vocaliques ou consonantiques et selon qu'elles sont typiquement masculines, féminines (donc marquées pour le genre grammatical) ou ambiguës (donc non marquées).

Nous pouvons remarquer que, pour certaines terminaisons, il y a plusieurs noms correspondants. S'il y a parfois un déséquilibre (*-eur* : sept noms), c'est que nous n'avons pas trouvé de noms substitués, mais avec d'autres terminaisons, pour la catégorie en question. Il est bon de noter aussi qu'il y a plus de noms à terminaisons non marquées (32 noms) que de noms à terminaisons marquées (28 noms, 16 masculines et 12 féminines).

De plus, chacun des noms a été étiqueté pour sa fréquence, telle que calculée dans le corpus de Baudot (1992), comme mentionné plus haut. Dans le Tableau 2.3, nous donnons la liste exhaustive des 60 noms sélectionnés pour nos tests avec cette fréquence.

Tableau 2.2
La répartition des terminaisons des 60 noms (vrais mots)

Terminaisons vocaliques	
Masculines (5 mots) (marquées)	-ent, -er, -é après une-C ($C \neq t$), -et, -ot
Féminines (8 mots) (marquées)	-ée (2), -ion (2), -ie, -oue, -té (2)
Ambiguës (9 mots) (non marquées)	-a (2), -son (3), -ia (2), -oix, -çon
Terminaisons consonantiques	
Masculines (11 mots) (marquées)	-isme (2), -age (3), -ège (2), -ort, -ème, -cl, -ar
Féminines (4 mots) (marquées)	-ette (2), -ille, -uche
Ambiguës (23 mots) (non marquées)	-ole, -oupe, -ange, -ire, -eur (7), -re après V(C)C ($C \neq l$) (3), -ère, -este, -ode, -ale, -ique, -agne, -ide, -upe, -oire

Pour déterminer si un nom français et son équivalent en espagnol étaient congénères, nous avons fait une première sélection qui était assez conservatrice, c'est-à-dire que nous avons éliminé les noms qui nous semblaient ambigus pour ce qui est de leur congénérarité. Ensuite, nous nous en sommes tenu à un accord interjuges. Pour ce faire, nous avons présenté les items que nous avions choisis à deux locuteurs natifs de l'espagnol (les mêmes qui ont vérifié les équivalents espagnols des noms français) et à deux locuteurs natifs du français. Si trois personnes sur quatre considéraient aisément que deux noms présentés étaient congénères, ces derniers étaient potentiellement sélectionnés pour faire partie de la liste définitive. Les congénères devaient être uniquement de vrais amis, donc des noms qui étaient similaires au niveau de la forme et qui possédaient le même sens, et ce, dans toutes leurs acceptions ou presque. Pour ce faire, nous avons d'abord présenté aux juges environ 500 noms, puis une centaine, une fois la liste épurée.

Tableau 2.3

Les 60 noms choisis (vrais mots) avec leur fréquence dans le corpus de Baudot

Mots	Fréquence				
		poésie	46	gorgée	5
moment	604	symbole	43	ferveur	4
famille	388	miel	36	fourchette	4
groupe	387	blé	30	jardinage	4
mot	371	piège	27	liège	4
couleur	254	pistolet	25	rançon	4
période	208	brique	23	gala	3
fleur	159	guérison	21	journalisme	3
peur	155	beurre	19	losange	3
volonté	154	montre	19	blasphème	2
choix	152	cauchemar	16	cafétéria	2
danger	117	punition	16	cidre	2
montagne	96	frisson	13	clarinette	2
navire	89	lâcheté	13	conifère	2
transport	83	laideur	12	panorama	2
ruche	53	tracteur	11	plumage	2
joue	52	nettoyage	10	nageoire	1
malheur	52	céréale	8	tibia	1
mécanisme	50	buisson	7	zeste	1
vallée	48	pyramide	6		
jupe	47	contagion	5		

Nous avons dû aussi tenir compte d'une répartition adéquate des terminaisons (terminaisons différentes le plus possible). Nous nous sommes assuré que les terminaisons à finales vocaliques et consonantiques étaient réparties de façon homogène. De plus, les fréquences des noms dans Baudot (1992) ont été distribuées de la façon la plus équitable possible (comme nous l'avons décrit auparavant).

Ce choix de mots a fourni un équilibre quasi binaire pour chaque catégorie. Cependant, il nous a fallu créer un léger déséquilibre du côté des noms féminins. En effet, il nous a été impossible de trouver tous les noms nécessaires pour les noms féminins à terminaisons marquées, mais ne se terminant pas en *-e*. Ce regroupement correspond aux terminaisons *-ion* et *-té*, très fréquentes en français. Comme il nous a semblé impossible de trouver les noms qui correspondaient à une combinaison de critère parfaite (après plusieurs heures de recherche et de vérification auprès d'hispanophones), nous avons omis quatre noms de la grille binaire. De plus, nous avons déjà quatre noms à terminaisons en *-ion* et en *-té*. Le total des noms aurait dû être de 64 (32 noms pour chaque genre), mais il est passé à 60 (32 noms masculins et 28 noms féminins).

2.5 La passation des tests

Une séance avec un participant se passait de la façon suivante :

D'abord, dans un local contenant un ordinateur portable, l'expérience se déroulait individuellement. Le participant devait lire et signer un formulaire de consentement; il remplissait un petit questionnaire de renseignements généraux afin de nous assurer d'un certain contrôle des variables sociobiographiques, puis il recevait quelques consignes générales. L'expérience comportait deux tests : d'abord le test de reconnaissance, ensuite le test principal. Une consigne et quelques numéros de pratique étaient administrés pour chacun des tests (voir Annexe B, pour les différents documents). Ces deux tests ont été créés et réalisés à l'aide du logiciel LÉA, confectionné par Michel Bastien, étudiant au doctorat en linguistique, à l'UQAM. L'utilisation du logiciel LÉA a permis :

- 1) une compilation facile des résultats;
- 2) une durée maximale de réflexion pour chaque item (dix secondes), qui était égale d'un sujet à l'autre;
- 3) la possibilité que l'apprenant soit soumis à un seul mot à la fois, et ce, sans retour en arrière possible;

- 4) une présentation aléatoire des items, ce qui évitait la contamination d'un item à l'autre;
- 5) l'ajout d'un masque (#####) entre les items, afin d'effacer le mot précédent sur la rétine de l'œil, au cas où le genre ou la forme de ce mot aurait influencé l'item en cours;
- 6) le calcul du temps de réponse, qui, bien que non essentiel dans notre recherche, pouvait nous donner certaines indications quant à l'hésitation potentielle d'un sujet lors d'une assignation.

Le premier test, le test de reconnaissance, était constitué des mêmes items que pour le test principal et était construit exactement de la même manière. Cependant, le participant devait appuyer sur la touche O pour « Oui » ou N pour « Non » du clavier, selon qu'il croyait connaître le sens du nom présenté à l'écran. Cette connaissance du sens d'un nom se voulait approximative : une seule acception suffisait. Ce test nous permettait de classer les mots (quantité de participants connaissant chacun des mots) comme les participants (quantité de mots connus par chacun des participants), et d'éventuellement nous aider à répondre à nos questions de recherche de façon plus précise. Si le participant ne répondait pas dans un temps limite de dix secondes, nous ne tenions pas compte de cette réponse.

Dans le deuxième test, le test principal, les participants devaient donc attribuer le genre à une liste de substantifs apparaissant à l'écran, en appuyant sur la touche M pour le masculin et la touche F pour le féminin. Le temps maximal de réponse pour chacun des items était de dix secondes (programmées dans LÉA). Comme pour le test de reconnaissance des noms, si un participant n'avait pas donné de réponse pour un item à l'intérieur des dix secondes allouées, LÉA passait automatiquement à l'item suivant et l'item n'était pas compté dans les résultats (codé « non répondu »).

La passation totale des tests par participant (incluant les formulaires à remplir et les explications) a été d'une durée d'environ 30 minutes (test de reconnaissance + test principal). Des pré-tests avaient été prévus au préalable afin de vérifier nos outils. Dix personnes, principalement francophones, ont été testées, et ce, au cours de la même semaine (début avril

2007). Ces pré-tests nous ont permis de nous pratiquer à les faire passer de la façon la plus identique possible pour tous les participants. Nous n'avons rien changé aux tests après ces pré-tests : seul le déroulement de leur passation a été affiné.

Quant aux vrais tests, le recrutement et la passation de ceux-ci nous ont pris environ deux mois et demi (fin avril, mai et juin 2007). Il y a eu des moments d'affluence et des moments d'accalmie : les tests ne se sont donc pas déroulés de manière homogène, quotidiennement ou hebdomadairement. La collaboration irrégulière des participants et le taux d'absentéisme expliquent cette durée prolongée. Il a souvent fallu reprendre rendez-vous avec certains participants qui ne se présentaient pas.

Les participants nous ont rarement posé des questions inattendues : les pré-tests nous avaient permis de bien nous préparer à ces éventuelles questions. Nous n'avons pas mentionné aux participants que les tests comportaient des non-mots. Cependant, les participants nous ont parfois demandé s'il pouvait y avoir des mots inexistant en français (les non-mots). Nous leur avons répondu que tous les mots étaient français et qu'on ne devait pas poser de question pendant les tests puisqu'ils n'avaient que dix secondes pour répondre à chaque item. Par contre, ces commentaires ont été dits avant le début des vrais tests. Une pratique de dix items, incluant aussi des non-mots, avait été prévue avant chacun des tests. Cela permettait de rassurer les participants et de soulever leurs questions avant le début des vrais tests. Pour les participants qui nous avaient posé des questions concernant l'existence de non-mots, nous leur avons dévoilé la vérité à la fin de la passation de tests. Nous pouvons affirmer que le comportement de l'ensemble des participants, hispanophones ou sinophones, était similaire. Cependant, surtout lorsque les participants étaient moins avancés en français, ils pouvaient être nerveux et se comporter comme si on les évaluait. Nous les avons rassurés au préalable que ces tests ne comptaient pas dans leur cursus universitaire et que toute réponse était valable. Il nous a semblé, tout de même, que ce cas de figure se retrouvait plutôt chez les sinophones.

2.6 Les procédés de compilation et d'analyse des résultats

Afin de bien répondre à nos questions de recherche, nous avons, d'une part, observé les résultats obtenus sous deux angles : les mots et les participants. Puis, d'autre part, nous avons

observé les résultats du test de reconnaissance des mots et du test d'assignation du genre des mots. Les non-mots ont subi un traitement spécial. Pour le test de reconnaissance, le total des non-mots répondus O (Oui), pour chacun des participants, a été soustrait du total des noms connus. Pour le test d'assignation du genre, les non-mots ont été observés séparément pour les résultats quant aux terminaisons. Nous rappelons que les items n'ayant pas reçu de réponse à l'intérieur de dix secondes ont été traités comme des non-réponses et n'ont pas fait partie des résultats.

Pour le test d'assignation du genre, trois analyses séparées ont été effectuées. La première était une comparaison des groupes linguistiques pour la paire de variables « Même genre » (MG) et Congénère » (C). La deuxième analyse reflétait les résultats pour les terminaisons en -e (e). La troisième correspondait aux terminaisons marquées ou non par le genre (m).

Finalement, toutes les analyses, que ce soit celles correspondant au test de reconnaissance ou au test d'assignation du genre, ont été mises en relation avec la variable « sexe des participants ». Les résultats globaux ont aussi été explorés : les taux d'erreurs et de réussites pour chaque groupe linguistique et pour chaque genre, et ce, indépendamment des variables précédemment mentionnées.

M. Bertrand Fournier, du Service de consultation en analyses de données de l'UQAM (SCAD) a procédé à des analyses statistiques utilisant le test de Pearson, le test de coefficient de corrélation de Kendall, le test bilatéral de Wilcoxon utilisant l'approximation normale, ainsi que le test bilatéral des rangs signés.

2.7 Conclusion

En somme, la méthodologie nous aura permis de répondre à nos différentes questions de recherche. Nous verrons, dans le prochain chapitre, si nos hypothèses ont été confirmées ou infirmées.

CHAPITRE III

PRÉSENTATION ET ANALYSE DES RÉSULTATS

3.1 Introduction

Dans ce chapitre, nous présenterons et analyserons les résultats des tests statistiques. Comme nous l'avons mentionné dans la méthodologie, ces tests ont été effectués par M. Bertrand Fournier, du Service de consultation en analyses de données de l'UQAM (SCAD). Nous diviserons notamment ce chapitre selon les neuf questions de recherche qui ont été posées au chapitre I. De plus, nous confirmerons ou infirmerons nos hypothèses sous chaque réponse à nos questions de recherche.

Nous voulons ici rappeler un aspect théorique qui concernera tous les résultats. Chaque fois que nous parlerons d'un recours à un processus ou à une stratégie (que ce soit en parlant du transfert, des terminaisons, ou d'autres choses), nous sommes conscient que cela ne constituera qu'une inférence, de ce qui ne pourrait, selon nous, être affirmé directement. En effet, nous ne pouvons parler que des résultats obtenus par les participants lors des tests, et non de ce qui se passe dans le cerveau des participants. Il faudrait souligner ici que nous ne pouvons jamais affirmer quelque chose avec certitude concernant ce qui se passe cognitivement parlant chez l'apprenant. Nous voulons donc ici avertir le lecteur que nous sommes conscient de cela et que, lorsque nous parlerons de transfert ou de recours à la stratégie du transfert (idem pour le recours aux terminaisons), nous ne ferons que suggérer la possibilité de ces processus cognitifs; nous émettrons, en quelque sorte, de nouvelles hypothèses d'ordre psycholinguistique, mais non confirmées, vérifiées dans cette recherche.

Finalement, nous voulons ici mentionner que nous avons pris les seuils de signification tels que présentés dans les analyses fournies par le SCAD. Nous sommes conscient que le

seuil acceptable est de $p < 0,05$, mais nous avons toujours déclaré les seuils de significations exacts, qui étaient parfois beaucoup moins élevés (ex. : $p < 0,0001$).

3.2 La réponse à la question de recherche générale

Q1 : De façon globale, quel groupe linguistique aura plus de succès dans l'assignation correcte du genre en français, celui dont la L1 possède un genre, les hispanophones, ou celui dont la L1 n'en possède pas, les sinophones?

D'après les résultats statistiques, les hispanophones ont des résultats d'assignation du genre grammatical très légèrement supérieurs aux sinophones, mais l'écart entre ces deux résultats moyens n'est pas significatif. En effet, les pourcentages moyens de réussite au test d'assignation du genre sont de 62,7 % pour les hispanophones (écart type de 6,9) et de 60,2 % pour les sinophones (écart type de 5,3). Pour ce test, le coefficient de corrélation de Kendall entre les deux groupes est de 0,15669, et le taux de probabilité est de 0,0872, ce qui est supérieur au seuil de signification de 0,05. Donc, les différences entre les hispanophones et les sinophones, pour l'assignation correcte du genre, ne sont pas significatives. Notons que, pour les hispanophones, le résultat le plus bas est de 51,7 %, le résultat le plus haut est de 80 %, et la médiane est de 62,5 %. Pour les sinophones, le résultat le plus bas est de 48,3 %, le résultat le plus haut est de 68,3 %, et la médiane est de 60,9 %.

La réponse à notre question de recherche principale et donc la suivante : les hispanophones et les sinophones ont des résultats globaux similaires lors de l'assignation du genre grammatical en français. Le fait de posséder ou non le genre grammatical dans sa langue maternelle ne semble pas influencer sur les résultats globaux d'assignation du genre grammatical, du moins, pour ces deux groupes linguistiques spécifiques. Mais, comme nous allons le voir, il existe néanmoins des différences importantes et intéressantes dans les comportements des deux groupes d'apprenants.

3.3 Les réponses aux questions de recherche sur le transfert

Nous allons maintenant répondre à nos deux questions de recherche concernant le recours au transfert de la L1. Nous voulons rappeler ici que le recours au transfert est une des deux catégories de stratégies potentiellement employées (selon notre supposition),

consciemment ou non, par les apprenants d'une langue seconde (l'autre étant le recours aux terminaisons des noms).

3.3.1 Les hispanophones

Q2 : Dans quelle mesure y a-t-il recours au transfert par les hispanophones?

Les résultats des hispanophones sont très éloquentes et nous font suggérer qu'il y aurait eu fortement recours au transfert.

3.3.2 Les congénères

Q3 : Quel est l'impact, positif ou négatif, des congénères sur ce transfert?

C'est l'effet des congénères qui a eu le plus grand impact sur l'assignation du genre grammatical par les hispanophones, et ce, de manière autant positive que négative.

Tableau 3.1
Pourcentage moyen de réussite (avec écart type) au test d'assignation du genre selon « même genre » et « congénère »

	C				NC			
	MG (15)		NMG (15)		MG (15)		NMG (15)	
	moyenne	écart type	moyenne	écart type	moyenne	écart type	moyenne	écart type
H	96,0	5,2	33,7	11,6	63,4	9,3	58,0	16,2
S	64,9	9,7	61,0	9,4	52,4	8,6	62,4	9,9

Dans le Tableau 3.1, on voit clairement que, lorsque les noms sont congénères (C), les hispanophones assignent principalement le genre de leur langue. En effet, quand les noms congénères avec l'espagnol ont le même genre qu'en espagnol (MG_C), le pourcentage moyen de réussite chez les hispanophones est très élevé (96,0 %). Même si nous avons discuté, dans la partie théorique de ce mémoire, qu'il n'était pas possible d'être certain de l'interprétation à donner lors de résultats positifs, en ce qui concerne le transfert, un taux de

réussite aussi élevé nous permet quand même de suggérer fortement qu'il y a ici un cas de transfert positif. Inversement, quand les congénères n'ont pas le même genre qu'en espagnol (NMG_C), ce pourcentage moyen de réussite est très faible (33,7 %). C'est donc un cas clair de transfert négatif. Cependant, pour ce qui est des noms non congénères avec l'espagnol (NC), l'écart entre les résultats MG et NMG est moins grand : 63,4 % de moyenne pour les MG_NC et 58,0 % de moyenne pour les NMG_NC. Ces derniers résultats montrent tout de même que, même si les noms sont non congénères, les résultats moyens sont supérieurs lorsque le nom possède le même genre dans les deux langues. Cette différence est significative ($p = 0,0487$, selon le test des rangs signés bilatéral). Cependant, toutes les autres comparaisons de résultats ont une $p < 0,0001$. En somme, cela démontre bien qu'il y a une tendance, chez les hispanophones, à se fier au genre de leur L1, mais massivement plus fortement quand ils reconnaissent clairement le nom, c'est-à-dire que celui-ci est congénère avec son homologue dans leur langue.

En ce qui concerne les résultats des sinophones, l'effet des congénères ne peut pas jouer puisque le français et le chinois ne possèdent pas de congénères entre eux. Rappelons que les libellés de MG (même genre) et de C (congénère) ont été choisis par rapport à notre groupe expérimental, le groupe des hispanophones. Le groupe des sinophones est, pour ce qui est des questions concernant le transfert, notre groupe contrôle. Les résultats des sinophones sont donc assez similaires selon la catégorie (voir Tableau 3.1). La différence entre les résultats de certaines catégories de noms est tout de même parfois significative (MG_C et MG_NC : $p < 0,0001$, NMG_C et MG_NC : $p = 0,0011$, MG_NC et NMG_NC : $p < 0,0001$). En bref, les résultats des MG_NC sont plus faibles que pour les autres catégories. Nous n'avons cependant pas d'explications à ce sujet. Finalement, nous pouvons globalement constater que chaque fois que nous avons la catégorie NMG, les sinophones ont de meilleurs résultats que les hispanophones. Cela est particulièrement vrai pour les NMG_C puisque, comme nous l'avons déjà dit, l'écart entre les résultats des MG et des NMG est accentué pour les hispanophones lorsque les noms sont congénères.

Selon le test bilatéral de Wilcoxon utilisant l'approximation normale (Tableau 3.2), nous pouvons comparer les résultats des hispanophones avec ceux des sinophones. Ce test répond à la question suivante : de combien l'écart type de l'échantillon est-il loin de la

moyenne? La cote z fournie par le test indique la distance et la direction de la déviance de cet échantillon, de sa moyenne de distribution, exprimées en unités de distribution des écarts types. Cette cote est donc un moyen de convertir la moyenne d'un certain ensemble de valeurs à 0 et leurs écarts types à 1. De cette façon, il est possible de normaliser et de comparer les données de deux ensembles de valeurs distincts en comparant leur cote z . Ici, si la cote z est supérieure à 0, c'est que les résultats des hispanophones sont supérieurs à ceux des sinophones. Inversement, si la cote z est négative, c'est que les résultats des hispanophones sont inférieurs à ceux des sinophones.

Tableau 3.2

Comparaison des pourcentages moyens de réussite des deux groupes linguistiques, avec le test bilatéral de Wilcoxon utilisant l'approximation normale, selon la combinaison des catégories « même genre » et « congénère » (test d'assignation du genre)

Catégorie des mots		z	$p > z $
C	MG (15)	6,6573	<0,0001
	NMG (15)	-6,1533	<0,0001
NC	MG (15)	4,3059	<0,0001
	NMG (15)	-1,3353	0,1818

Nous constatons que, pour les non-congénères, la différence est significative entre les deux groupes linguistiques lorsque les noms ont le même genre que pour l'espagnol ($z = 4,3059$, $p < 0,0001$). Nous suggérons une explication de ce phénomène par le choix des noms ou par l'effet d'autres facteurs que nous discuterons plus loin. En bref, même si les noms ne sont pas congénères, ils sont mieux réussis par les hispanophones que par les sinophones lorsqu'ils possèdent le même genre (en français comme en espagnol). Cependant, nous confirmons ici que, pour les congénères, l'écart entre les résultats moyens des sinophones et celui des hispanophones est très éloquent et significatif ($z = 6,6573$ et $z = -6,1533$, $p < 0,0001$ dans les deux cas). Nous suggérons l'explication suivante : il y a transfert positif (MG_C) et transfert négatif (NMG_C) dans le cas des résultats des hispanophones, ce qui fait contraste avec ceux des sinophones.

Nous pouvons donc dire que notre hypothèse selon laquelle les hispanophones auraient grandement recours au transfert est confirmée, dans la mesure où cela est suggéré par les

résultats. Cependant, nous pensions que les résultats pour les non-congénères auraient montré une plus grande différence entre les MG et les NMG. Nous croyions que cette différence aurait été significative et qu'indépendamment de la congénarité des noms, nous aurions pu clairement suggérer un effet du transfert. Cette conclusion montre clairement que l'effet des congénères est excessivement important lors de l'assignation du genre grammatical par des locuteurs d'une langue qui possède aussi des genres grammaticaux : l'hypothèse du recours au transfert du genre de la LI est très fortement accentuée. Lorsque les hispanophones assignent le genre à un nom qui est congénère avec un nom dans leur langue, ils assignent majoritairement le genre de leur langue, ce qui donne un transfert positif dans le cas des assignations réussies, et un transfert négatif dans le cas des assignations non réussies. Nous suggérons grandement que ce transfert négatif soit la source des résultats globaux, similaires à ceux des sinophones (nous rappelons que les résultats globaux au test d'assignation du genre sont les suivants : 62,7 % pour les hispanophones et de 60,2 % pour les sinophones). En effet, même si les hispanophones possèdent les mêmes genres qu'en français (mais distribués différemment) et qu'ils peuvent assez facilement reconnaître plusieurs noms français (les congénères), ils n'assignent pas globalement mieux le genre que les sinophones. Finalement, nous nous demandons si la fréquence des noms ou la connaissance du vocabulaire des participants (notamment celle des hispanophones, vu la grande quantité de congénères dans le corpus) ne joue pas sur les résultats. Nous verrons ces aspects dans la section 3.5.

3.4 Les réponses aux questions de recherche sur les terminaisons

Nous allons maintenant répondre aux deux questions de recherche concernant l'effet des terminaisons des noms, deuxième grande stratégie potentielle suggérée dans notre recherche.

Ici, nous rappelons au lecteur ce que nous entendons par terminaisons marquées pour le genre et par terminaisons en *-e*. Les terminaisons marquées pour le genre correspondent aux terminaisons typiquement féminines et typiquement masculines du corpus de Lyster (2006) (voir les Tableaux 1.4 et 1.5). Les terminaisons non marquées pour le genre correspondent aux terminaisons typiquement ambiguës (voir le Tableau 1.6). Par exemple, les terminaisons *-ion*, *-lé*, *-enne* et *-ette* sont typiquement féminines et les terminaisons *-age*,

-isme, *-eau* et *-ail* sont typiquement masculines. Ces huit terminaisons sont donc marquées pour le genre. Cependant, les terminaisons *-ia*, *-oix*, *-erbe* et *-agne* sont typiquement ambiguës, donc non marquées pour le genre. En ce qui concerne les terminaisons en *-e* muet final (toujours à l'écrit), leur nom décrit bien ce qu'elles sont. Parmi les terminaisons marquées et non marquées que nous venons de citer, *-ion*, *-té*, *-eau*, *-ail*, *-ia* et *-oix* sont des terminaisons sans *-e* muet final, tandis que *-enne*, *-ette*, *-age*, *-isme*, *-erbe* et *-agne* sont des terminaisons avec un *-e* muet final.

3.4.1 Les terminaisons marquées pour le genre

Q4 : Est-ce que les terminaisons marquées pour le genre (typiquement féminines et typiquement masculines) suscitent de meilleurs résultats que les terminaisons non marquées pour le genre (typiquement ambiguës)?

Il semble y avoir un clair effet des terminaisons marquées pour le genre sur l'assignation de celui-ci.

Tableau 3.3
Pourcentage moyen de réussite (avec écart type) au test d'assignation du genre selon « genre du nom » et « terminaison marquée »

	m				Nm			
	F (12)		M (16)		F (16)		M (16)	
	moyenne	écart type	moyenne	écart type	moyenne	écart type	moyenne	écart type
H	69,2	11,8	71,9	10,1	54,3	15,2	57,4	13,7
S	81,7	12,8	58,8	13,4	53,6	10,1	52,0	9,7

Pour les deux groupes linguistiques, le pourcentage moyen de réussite est supérieur lorsque les noms possèdent des terminaisons marquées pour le genre grammatical (voir le Tableau 3.3). Du côté des hispanophones, les moyennes sont respectivement de 69,2 % pour les noms féminins à terminaisons marquées pour le genre, et de 71,9 % pour les noms

masculins. Comparativement, les résultats des noms à terminaisons non marquées pour le genre sont de 54,3 % pour les noms féminins et de 57,4 % pour les noms masculins. Selon le test bilatéral des rangs signés, seules les différences entre les catégories à terminaisons marquées pour le genre et les catégories à terminaisons non marquées pour le genre sont significatives (F_m et F_Nm : $p < 0,0001$, F_m et M_Nm : $p = 0,0013$, M_m et F_Nm : $p < 0,0001$, M_m et M_Nm : $p < 0,0001$).

Pour ce qui est des sinophones, les moyennes sont de 81,7 % pour les noms féminins à terminaisons marquées pour le genre, et de 58,8 % pour les noms masculins. Les résultats des noms à terminaisons non marquées pour le genre, pour ce groupe linguistique, sont de 53,6 % pour les noms féminins et de 52,0 % pour les noms masculins. Nous constatons ici un très haut succès des sinophones pour la catégorie des terminaisons féminines marquées pour le genre. Cette catégorie est beaucoup mieux réussie par les sinophones que par les hispanophones. Cependant, pour la catégorie des noms masculins à terminaisons marquées pour le genre, ce sont les hispanophones qui ont obtenu de meilleurs résultats. Selon le test bilatéral des rangs signés, seules les différences entre la catégorie des noms féminins à terminaisons marquées pour le genre et toutes les autres catégories sont significatives (F_m et M_m : $p < 0,0001$, F_m et F_Nm : $p < 0,0001$, F_m et M_Nm : $p < 0,0001$).

Tableau 3.4

Comparaison des pourcentages moyens de réussite des deux groupes linguistiques, avec le test bilatéral de Wilcoxon utilisant l'approximation normale, selon la combinaison des catégories « genre du nom » et « terminaison marquée » (test d'assignation du genre)

Catégorie des mots		z	$p > z $
m	F (12)	-3,6509	0,0003
	M (16)	4,0577	<0,0001
Nm	F (16)	0,0225	0,9820
	M (16)	1,6167	0,1059

Nous constatons donc que, dans l'ensemble, et ce, pour les deux groupes linguistiques, les terminaisons marquées pour le genre ont de meilleurs taux d'assignation que les terminaisons non marquées. Cependant, pour les sinophones, le contraste entre les terminaisons marquées féminines et les terminaisons marquées masculines est plus prononcé.

La différence entre les résultats de la catégorie des terminaisons marquées masculines et des terminaisons non marquées n'est pas significative. Nous n'avons pas d'explications pour ce phénomène.

Selon le test bilatéral de Wilcoxon utilisant l'approximation normale (Tableau 3.4), nous pouvons encore comparer les résultats des hispanophones à ceux des sinophones. Nous constatons qu'il y a une différence significative entre les deux groupes linguistiques pour les noms à terminaisons marquées pour le genre grammatical ($p = 0,0003$ et $<0,0001$), mais pas pour les noms à terminaisons non marquées pour celui-ci ($p = 0,9820$ et $0,1059$). Encore une fois, nous n'avons pas d'explication claire pour ces différences. Nous pensons que ces différences s'expliquent probablement par le choix des mots.

3.4.2 Les terminaisons en *-e* muet

Q5 : Est-ce que les noms qui ont une terminaison en *-e* muet final, à l'écrit, (indépendamment du fait que leurs terminaisons soient marquées ou non pour le genre) incitent les apprenants à assigner le genre féminin et est-ce que les noms qui n'ont pas de terminaison en *-e* muet final incitent les sujets à assigner le genre masculin?

Pour ce qui est de l'effet des terminaisons en *-e* muet sur l'assignation du genre grammatical, les résultats parlent d'eux-mêmes (voir Tableau 3.5).

Pour ce qui est des hispanophones, les terminaisons en *-e* muet ne semblent pas influencer fortement leurs choix de réponses. On constate tout de même une amélioration de leurs résultats pour les noms qui n'ont pas une terminaison en *-e* muet. Selon le test bilatéral des rangs signés, les seules différences significatives sont celles entre les catégories des terminaisons en *-e* et la catégorie des noms masculins à terminaisons sans *-e* (F_e et M_Ne : $p = 0,0002$, M_e et M_Ne : $p = 0,0005$) Nous voyons un début d'explication de cette amélioration des résultats pour les noms masculins à terminaisons sans *-e* muet final dans le fait que les terminaisons sans *-e* prédisent plus facilement le genre masculin que les terminaisons avec un *-e* ne prédisent le genre féminin (voir Lyster, 2006, dans le chapitre 1).

Les résultats des sinophones, concernant les terminaisons en *-e*, sont, quant à eux, beaucoup plus parlant que leurs résultats selon les terminaisons marquées pour le genre ou

que ceux des hispanophones pour les terminaisons en *-e*. Nous suggérons que les sinophones assignent le féminin à un nom qui possède un *-e* final muet, et assignent le masculin à un nom qui n'en possède pas. Cela semble d'ailleurs être leur stratégie la plus probable.

Tableau 3.5
Pourcentage moyen de réussite (avec écart type) au test d'assignation du genre selon
« genre du nom » et « terminaison en *-e* »

	e				Ne			
	F (16)		M (16)		F (12)		M (16)	
	moyenne	écart type	moyenne	écart type	moyenne	écart type	moyenne	écart type
H	58,8	12,0	59,4	13,7	63,1	17,9	69,9	10,5
S	88,7	10,8	23,0	17,6	35,0	19,4	87,8	10,3

Les données sont claires : les taux de réussite sont très élevés pour les noms féminins qui se terminent en *-e* (88,7 %) et pour les noms masculins qui ne se terminent pas en *-e* (87,8 %), tandis que ces taux sont très faibles pour les noms masculins qui se terminent en *-e* (23,0 %) et pour les féminins qui ne se terminent pas en *-e* (35,0 %). Selon le test bilatéral des rangs signés, toutes les différences entre les catégories de noms sont très significatives ($p < 0,0001$ et $p = 0,0038$ pour la différence entre M_e et F_{Ne}). La seule différence qui n'est pas significative est celle entre F_e et M_{Ne} . Ces résultats s'expliquent facilement par les taux de réussite massivement supérieurs lorsque le genre des noms correspond à la présence ou à l'absence de *-e* muet final. De plus, la hausse du taux de réussite des noms féminins sans *-e* par rapport au taux de réussite des noms masculins avec un *-e* correspond probablement au haut taux de réussite dans la catégorie des noms féminins à terminaisons marquées pour le genre (voir le Tableau 3.3).

À nouveau, le test bilatéral de Wilcoxon utilisant l'approximation normale (Tableau 3.6), nous aide à comparer les résultats entre les hispanophones et les sinophones. Nous pouvons constater que la différence des pourcentages moyens de réussite entre les deux

groupes linguistique est très significative ($p < 0,0001$ partout). Cela conduit fermement à la conclusion que les sinophones semblent utiliser la stratégie d'assignation du genre grammatical selon la terminaison en *-e* ou non des noms. La stratégie de l'assignation du genre selon que la terminaison des noms est en *-e* muet final ou non est plus difficile à affirmer pour les hispanophones puisque l'effet de la stratégie du transfert de leur L1 est omniprésent.

Tableau 3.6

Comparaison des pourcentages moyens de réussite des deux groupes linguistiques, avec le test de Wilcoxon bilatéral utilisant l'approximation normale, selon la combinaison des catégories « genre du nom » et « terminaison en *-e* » (test d'assignation du genre)

Catégorie des mots		z	p > z
e	F (16)	-6,0831	<0,0001
	M (16)	5,9571	<0,0001
Ne	F (12)	4,7574	<0,0001
	M (16)	-5,2293	<0,0001

Comme nous le verrons à la section 3.6, les résultats obtenus avec les non-mots confirment cette différence entre les hispanophones et les sinophones nous rappelant que les sinophones ne peuvent pas se référer à leur langue (pas de noms connus).

En somme, nous confirmons dans l'ensemble nos hypothèses concernant l'effet des terminaisons. Les noms à terminaisons marquées pour le genre obtiennent de meilleurs résultats d'assignation du genre. Ces résultats sont cependant dans l'ombre des résultats d'assignation du genre selon que la terminaison des noms (marquée ou non pour un genre particulier) se termine en *-e* muet final ou non. La stratégie du recours aux terminaisons semble donc être efficiente, mais celle du recours aux terminaisons en *-e* muet final prime sur celle du recours aux terminaisons marquées pour le genre, surtout pour les sinophones, qui ne peuvent pas avoir accès à leur L1, dans le contexte. Cela confirme donc notre hypothèse selon laquelle l'effet des terminaisons en *-e* (ou non) aurait un plus gros impact sur l'attribution du genre que l'effet des terminaisons marquées (ou non) pour le genre. Cet impact des terminaisons en *-e* muet final demeure surtout présent pour les sinophones. Le fait

que les terminaisons des noms soient marquées ou non pour le genre a cependant un impact pour les deux groupes linguistiques.

3.5 Les réponses aux autres questions de recherche

Nous tenterons maintenant de répondre aux quatre dernières questions de recherche posées. Les deux prochaines questions nécessiteront les résultats du test de reconnaissance (Tableau 3.7). Nous rappelons que ce test s'est déroulé avant le test principal (d'assignation du genre), et ce, pour chaque participant. Ce test nous a permis, notamment, de classer les noms selon le nombre de participants qui disaient reconnaître chacun d'entre eux, et de classer les participants selon le nombre de noms que chacun disait reconnaître.

Tableau 3.7
Pourcentage moyen de réussite (avec écart type) au test de reconnaissance
selon « même genre » et « congénère »

	C				NC			
	MG (15)		NMG (15)		MG (15)		NMG (15)	
	moyenne	écart type	moyenne	écart type	moyenne	écart type	moyenne	écart type
H	99,8	1,3	88,3	7,6	54,3	16,1	79,9	11,4
S	72,9	16,8	41,6	17,4	42,2	17,2	62,9	17,3

Contrairement au test d'assignation du genre, dans lequel on ne pouvait pas déceler de différences majeures en ce qui concerne les résultats globaux des deux groupes linguistiques (voir section 3.2), le test de reconnaissance donne des résultats globaux significativement différents d'un groupe linguistique à l'autre. Du point de vue de la reconnaissance des noms affichés, les hispanophones ont une moyenne totale de 78,2 % (écart type de 7,9), tandis que les sinophones ont une moyenne totale de 52,7 % (écart type de 14,4). Pour le test de reconnaissance, le coefficient de corrélation de Kendall entre les deux groupes est de 0,57614, et le taux de probabilité est de $< 0,0001$, ce qui est très significatif. Notons que, pour les hispanophones, le résultat le plus bas est de 63,3 %, le résultat le plus haut est de 93,3 %,

et la médiane est de 76,7 %. Pour les sinophones, le résultat le plus bas est de 13,6 %, le résultat le plus haut est de 78,3 %, et la médiane est de 53,8 %.

Rappelons ici que le test de reconnaissance contenait des non-mots et que le nombre de non-mots dits reconnus a été soustrait du nombre total de noms dits reconnus³. Sans la réduction des scores par les non-mots, ces résultats sont de 80,6 %, pour les hispanophones, et de 54,9 %, pour les sinophones, ce qui est supérieur. Les résultats montrent clairement que les hispanophones disent reconnaître plus de noms que les sinophones. Bien sûr, cette différence entre les deux groupes s'explique, entre autres, par le fait que la moitié du corpus est constituée de congénères avec les noms espagnols. Pourtant, les résultats pour les non-congénères sont aussi supérieurs (mieux reconnus) pour les hispanophones.

Nous constatons que, pour les hispanophones, les résultats MG_NC (54,3 %) sont de beaucoup inférieurs aux résultats NMG_NC (79,9 %). Nous constatons aussi que les différences entre les catégories MG_C et NMG_C, ainsi que celles entre MG_NC et NMG_NC sont très prononcées, et ce, pour les deux groupes. Selon les résultats du test des rangs signés bilatéral, toutes les différences entre chaque catégorie sont extrêmement significatives ($p < 0,0001$ ou $p = 0,0001$) pour les hispanophones. En ce qui concerne les sinophones, toutes les différences entre chaque catégorie sont extrêmement significatives aussi ($p < 0,0001$), sauf pour une catégorie, celle entre NMG_C et MG_NC, qui est absolument nulle ($p = 1$). Nous n'avons pas d'explications pour ce dernier phénomène.

Pour les congénères, il y a une baisse du taux de reconnaissance lorsque les noms ne possèdent pas le même genre avec l'espagnol. Pour les non-congénères, c'est une hausse du taux de reconnaissance que nous constatons lorsque les noms n'ont pas le même genre avec l'espagnol. Nous ne possédons pas d'explications pour ces variations non négligeables à part l'effet arbitraire des noms choisis. Cependant, les résultats de la catégorie des noms qui n'ont pas le même genre qu'en espagnol et qui sont congénères sont explicables par le fait que les hispanophones reconnaissent les congénères. Nous pouvons donc affirmer que, ne tenant pas

compte de la différence due à la reconnaissance des congénères par les hispanophones, les résultats suivent sensiblement la même tendance pour les deux groupes linguistiques, mais les résultats des hispanophones sont supérieurs. Rappelons ici que les libellés MG et C se rapportent aux hispanophones, mais que les ensembles de noms représentés par ces libellés donnent des résultats observables pour notre groupe contrôle, les sinophones.

Tableau 3.8

Comparaison des pourcentages moyens de réussite des deux groupes linguistiques, avec le test bilatéral de Wilcoxon utilisant l'approximation normale, selon la combinaison des catégories « même genre » et « congénère ». (test de reconnaissance)

Catégorie des mots		z	p > z
C	MG (15)	6,7574	<0,0001
	NMG (15)	6,6362	<0,0001
NC	MG (15)	2,7087	0,0068
	NMG (15)	4,0352	<0,0001

Le test bilatéral de Wilcoxon utilisant l'approximation normale (Tableau 3.8), nous montre que la différence de résultats entre les deux groupes linguistiques est très significative ($p < 0,0001$). Elle l'est aussi pour les noms qui ont le même genre qu'en espagnol et qui ne sont pas congénères avec cette dernière langue ($p = 0,0068$). Comme nous l'avons dit auparavant, les différences de taux de reconnaissance des noms les plus probantes sont celles des noms qui sont congénères entre le français et l'espagnol ($z = 6,7574$ et $6,6362$).

3.5.1 La fréquence des noms

Q6 : Le taux d'assignation correcte du genre est-il fonction de la fréquence des noms présentés?

Ici, nous allons d'abord rappeler ce que nous entendons par *fréquence*. Nous avons eu deux façons d'obtenir la fréquence des noms. D'abord, les noms ont tous été choisis dans le

³ Il est à noter ici que, dans les résultats les plus bas, les résultats les plus hauts et la médiane, nous n'avons pas tenu compte de la soustraction des non-mots. Nous en avons tenu compte que pour la moyenne et l'écart type.

corpus de Baudot (1992); corpus, rappelons-le, qui donne un taux de fréquence d'emploi pour chaque item lexical retenu. Nous connaissions donc le taux de fréquence de chacun des mots que nous avons retenus pour le test de reconnaissance puisqu'ils ont tous été tirés du corpus de Baudot (1992). Ensuite, les participants ont déterminé, par le test de reconnaissance, s'ils reconnaissaient ou non les noms. Nous allons donc d'abord vérifier s'il y a corrélation entre la fréquence des noms selon le corpus Baudot et le taux de reconnaissance de ces noms par les participants, lors du test de reconnaissance.

Le coefficient de corrélation de Kendall entre le taux de reconnaissance des noms par les hispanophones et la fréquence des noms selon le corpus Baudot est de 0,31112, avec un taux de probabilité de 0,0012. Pour les sinophones, ce même coefficient de corrélation de Kendall est de 0,49069, avec un taux de probabilité de $< 0,0001$. Ces résultats sont donc très significatifs, et ce, pour les deux groupes linguistiques. Cela nous amène donc à affirmer que la fréquence d'emploi ou d'occurrence des noms choisis dans le corpus Baudot correspond assez fidèlement aux résultats de reconnaissance des noms par nos participants hispanophones et sinophones. Plus un nom du test de reconnaissance a un taux de fréquence élevé selon le corpus Baudot, plus le taux de reconnaissance de ce mot est élevé chez nos participants, tant chez les hispanophones que chez les sinophones. Nous pouvons dès lors parler de « fréquences / reconnaissances des noms », autant pour faire référence aux fréquences d'emploi selon le corpus Baudot que pour le taux de reconnaissance des noms lors du test de reconnaissance.

Nous pouvons maintenant répondre à notre prochaine question de recherche. Voyons premièrement s'il y a corrélation entre le taux de réussite de chacun des deux groupes linguistiques, lors de l'assignation du genre grammatical (test d'assignation du genre), et la fréquence des noms selon le corpus Baudot. Pour ces données, le coefficient de corrélation de Kendall chez les hispanophones est de 0,25672, avec un taux de probabilité de 0,0048, ce qui est très significatif. Chez les sinophones, le coefficient est de 0,17986, avec un taux de probabilité de 0,0480, ce qui est aussi significatif. Il est à noter que, bien que les résultats pour les deux groupes linguistiques sont significatifs, les coefficients de corrélation sont assez faibles.

Selon ces résultats, on observe que, pour les hispanophones, plus un mot est fréquent, plus il y a de chance qu'on lui assigne le bon genre. Pour les sinophones, il semble y avoir aussi un lien entre la bonne assignation du genre grammatical et la fréquence des noms observés.

Nous avons observé les taux de réussite (test d'assignation du genre) de l'ensemble des 28 hispanophones pour chacun des 30 noms qui étaient non congénères avec les noms espagnols. De cette façon, nous avons voulu vérifier si les taux de réussite d'assignation du genre étaient corrélés à la fréquence des noms dans Baudot lorsqu'on enlevait l'effet des congénères sur la reconnaissance des noms par les hispanophones (se rappelant que la fréquence des noms, dans Baudot, et la reconnaissance des noms lors du test de reconnaissance sont significativement corrélées). La moyenne des fréquences des noms non congénères est de 44,63 (1 étant le moins fréquent (*nageoire* et *zeste*) et 371 le plus fréquent (mot), pour les non-congénères). La moyenne des taux de réussite d'assignation du genre pour chacun de ces mêmes noms est de 60,77 % (14 % étant le résultat le plus faible (*brique*) et 100 % le résultat le plus élevé (*danger*)).

Neuf noms sur trente possèdent une fréquence supérieure à la moyenne. Parmi ces neuf noms, sept dénotent un taux de réussite supérieur à la moyenne. Ces résultats pourraient appuyer l'hypothèse selon laquelle la fréquence des noms et le taux de réussite d'assignation du genre de ceux-ci sont corrélés. Par contre, nous avons aussi observé qu'il y avait neuf autres noms, cette fois-ci à fréquence sous la moyenne, pour lesquels le taux de réussite d'assignation du genre était supérieur à la moyenne, ce qui diminue la force de ces résultats. Finalement, il reste tout de même douze noms non congénères à fréquence sous la moyenne, qui dénotent un taux de réussite inférieur à la moyenne.

Notre hypothèse selon laquelle plus un nom est fréquent plus son genre grammatical sera connu est donc confirmée, mais faiblement. Nous avons constaté que les résultats sont significatifs pour les sinophones comme pour les hispanophones.

Il est intéressant de constater que le mot le plus fréquent de notre corpus (*moment*, avec une fréquence de 604 dans le corpus de Baudot) a été reconnu à 100 % par les hispanophones et à 94 % par les sinophones. Le taux de réussite d'assignation du genre, pour ce mot, est de

100 % pour les hispanophones et de 97 % pour les sinophones. Les résultats de ce seul mot viennent confirmer ce que nous avons dit concernant la corrélation qui existe entre les fréquences dans Baudot et la reconnaissance des noms lors de notre test de reconnaissance. Pour ce mot très fréquent et très reconnu par les participants, les résultats d'assignation du genre sont aussi très élevés, et ce, pour les deux groupes linguistiques.

3.5.2 Le niveau de connaissance du vocabulaire des participants

Q7 : Le taux d'assignation correcte du genre est-il fonction du niveau de connaissance du vocabulaire des apprenants?

Maintenant que nous avons discuté les corrélations entre la fréquence des noms et l'assignation du genre grammatical, observons celles qu'il y a entre le niveau de connaissance des mots stimuli des participants et l'assignation correcte du genre grammatical. Nous rappelons que, dans notre recherche, le niveau de connaissance des mots stimuli correspond aux résultats au test de reconnaissance (voir Tableau 3.7).

Il ne s'agit pas d'évaluer le niveau global des compétences en français des participants (ni même en vocabulaire général), mais seulement leur niveau correspondant aux mots stimuli. De plus, nous ne parlons pas d'un niveau de connaissance des mots stimuli prédéterminé, mais bien des résultats basés sur leurs dires, et concernant les seuls noms de notre test. En d'autres termes, pour les besoins de notre recherche, plus un participant indique qu'il « reconnaît » de mots, plus nous considérons qu'il « connaît » de mots stimuli et plus nous le considérons « avancé » en connaissance du vocabulaire, en général (ici, nous sommes conscient que nous faisons une inférence quant à la connaissance générale du vocabulaire des participants).

Les coefficients de corrélation de Kendall, pour chacun des groupes linguistiques, entre les résultats de reconnaissance des noms (test de reconnaissance) et ceux d'assignation correcte du genre grammatical (test d'assignation du genre) sont les suivants : pour les hispanophones, le coefficient est de 0,27561, avec un taux de probabilité de 0,0045; pour les sinophones, le coefficient est de 0,11254, avec un taux de probabilité de 0,2189.

Les résultats de corrélation des hispanophones sont donc très significatifs, mais ils ne le sont pas pour les sinophones. L'effet du degré de connaissance du vocabulaire de la part du participant sur l'assignation correcte du genre semble donc réel pour les hispanophones, mais pas pour les sinophones, ce qui veut dire que pour ceux-ci, le fait de « reconnaître » un mot n'implique pas qu'on en connaît le genre approprié. Encore une fois, nous voyons l'effet des congénères dans ces résultats. Nous ne pouvons toutefois pas affirmer que le niveau de connaissance du vocabulaire d'un apprenant allophone a un lien univoque avec la bonne assignation du genre grammatical par celui-ci.

Cette fois-ci, nous avons observé les taux de réussite (test d'assignation du genre) de chacun des 28 hispanophones pour l'ensemble des 30 noms qui étaient non congénères avec les noms espagnols. Nous les avons comparés aux taux de reconnaissance de ces mêmes noms et pour ces mêmes participants. Il faut noter que les catégories MG_NC et NMG_NC ont été comptées séparément. Nous avons voulu vérifier si les taux de réussite d'assignation du genre étaient corrélés à la reconnaissance des noms lorsqu'on enlevait l'effet des congénères sur la reconnaissance des noms par les hispanophones.

La moyenne des taux de reconnaissance des noms non congénères est de 67,1 % (33,3 % étant le résultat le plus faible et 100 % le résultat le plus élevé). La moyenne des taux de réussite d'assignation du genre pour ces mêmes noms est de 60,7 % (26,7 % étant le résultat le plus faible et 100 % le résultat le plus élevé):

Vingt-sept taux de reconnaissance sur cinquante-six (ces taux peuvent faire partie des MG ou des NMG, fois 28 hispanophones) sont supérieurs à la moyenne. Parmi ces vingt-sept taux de reconnaissance, seulement onze catégories dénotent un taux de réussite d'assignation du genre supérieur à la moyenne. Ces résultats ne peuvent pas appuyer l'hypothèse selon laquelle la connaissance des noms et le taux de réussite d'assignation du genre de ceux-ci sont corrélés.

Notre hypothèse selon laquelle les apprenants « avancés en connaissance du vocabulaire français » auront de meilleurs résultats lors de l'assignation du genre grammatical (en français) est donc infirmée.

3.5.3 L'hypothèse du masculin par défaut

Q8 : De façon générale, est-ce que les sujets ont tendance à assigner le masculin par défaut (genre non marqué)?

Pour savoir si les participants ont assigné le masculin par défaut, nous ne pouvons pas observer les taux de réussite, il nous faut observer les erreurs des participants lors du test d'assignation du genre. À la lumière de la distribution des erreurs des participants à notre recherche (voir Tableaux 3.9 et 3.10), le genre masculin n'est pas assigné par défaut. En fait, il y a plus d'erreurs de noms masculins, auxquels les participants ont assigné le genre féminin, et ce, pour les deux groupes linguistiques. Nous rappelons que les résultats de la distribution des erreurs sont basés sur les 28 noms féminins (46,67 %) et 32 noms masculins (53,33 %) qui forment le corpus de nos tests (60 noms, excluant les non-mots). De plus, nous avons testé 28 hispanophones et 32 sinophones, ce qui explique, notamment, un plus gros échantillon de résultats pour les sinophones.

Tableau 3.9
Distribution des erreurs selon le genre grammatical pour le groupe des hispanophones

Vrai genre	Assignations erronées	Pourcentage d'erreurs	Pourcentage théorique
F (28 items)	307	49,20 %	46,67 %
M (32 items)	317	50,80 %	53,33 %

Test du Khi-2 # pour proportions spécifiées

Khi-2 1,6032
DDL 1
P > 0,2055

Taille de l'échantillon = 624

Dans le Tableau 3.9, on observe la distribution des erreurs des hispanophones. L'échantillon total des réponses erronées pour tous les participants de ce groupe linguistique et pour tous les mots est de 624 occurrences d'erreurs sur 1 676 occurrences totales (erreurs et réussites de tous les items auxquels les 28 hispanophones ont répondu). Ce total est divisé

en 307 occurrences (49,20 %) de noms féminins pour lesquels les hispanophones ont assigné le genre masculin (sur 780 occurrences totales de noms féminins), et en 317 occurrences (50,80 %) de noms masculins pour lesquels les hispanophones ont assigné le genre féminin (sur 896 occurrences totales de noms masculins). Le test de Pearson donne un Khi-2 de 1,6032, avec un taux de probabilité de 0,2055, ce qui n'est pas significatif comme différence de résultats d'erreurs par genre grammatical.

Tableau 3.10
Distribution des erreurs selon le genre grammatical pour le groupe des sinophones

Vrai genre	Assignations erronées	Pourcentage d'erreurs	Pourcentage théorique
F (28 items)	307	40,24 %	46,67 %
M (32 items)	456	59,76 %	53,33 %

Test du Khi-2 # pour proportions spécifiées

Khi-2	12,6908
DDL	1
P >	0,0004

Taille de l'échantillon = 763

Dans le Tableau 3.10, on observe la distribution des erreurs des sinophones. L'échantillon total des réponses erronées pour tous les participants de ce groupe linguistique et pour tous les mots est de 763 occurrences d'erreurs sur 1 917 occurrences (erreurs et réussites de tous les items auxquels les 32 sinophones ont répondu). Ce total est divisé en 307 occurrences (40,24 %) de noms féminins pour lesquels les sinophones ont assigné le genre masculin (sur 894 occurrences totales de noms féminins), et en 456 occurrences (59,76 %) de noms masculins pour lesquels les sinophones ont assigné le genre féminin (sur 1 023 occurrences totales de noms masculins). Le test de Pearson donne un Khi-2 de 12,6908, avec un taux de probabilité de 0,0004, ce qui est très significatif comme différence de résultats d'erreurs par genre grammatical.

Notre hypothèse selon laquelle les erreurs des participants aux tests devraient tendre vers le masculin est donc infirmée. C'est le contraire qui se produit, du moins pour les sinophones. En effet, les résultats significatifs des sinophones montrent qu'ils assignent plus souvent le féminin, lorsqu'ils devraient assigner le masculin. Notre explication, à ce sujet, est que huit noms sur les 60 noms choisis pour les tests sont des noms masculins, mais se terminant en *-e* muet final (mécanisme, plumage, liège, journalisme, blasphème, jardinage, nettoyage et piège). Comme nous l'avons vu auparavant, les sinophones ont clairement tendance à assigner le genre féminin aux noms se terminant en *-e* muet final. Notre hypothèse selon laquelle le masculin serait assigné par défaut ne peut pas être confirmée parce que les tendances d'assignation des hispanophones, comme des sinophones, sont très fortes : les hispanophones utilisent le genre de leur langue, surtout lorsqu'ils reconnaissent les noms, et les sinophones attribuent un genre selon que le nom possède ou non un *-e* muet final.

3.5.4 La variable « sexe »

Q9 : Y a-t-il un lien entre le sexe des participants et leurs résultats (corrélations entre le sexe et l'assignation du genre, meilleurs résultats pour un sexe donné)?

Comme nous l'avons mentionné au préalable, nous avons, de façon secondaire, observé les corrélations possibles entre les divers résultats et le sexe des participants. Les résultats concernant ces corrélations ne sont pas significatifs (tous les $p > 0,05$). Le sexe des apprenants ne joue pas un rôle dans l'assignation du genre grammatical, et ce, autant pour les hispanophones que pour les sinophones. Il n'y a pas de lien non plus entre le sexe des apprenants et leur reconnaissance des noms français. Notre hypothèse selon laquelle un homme aurait eu tendance à assigner de préférence le masculin et une femme, de préférence le féminin, est donc infirmée. De plus, comme nous le pensions, il n'y a pas un sexe en particulier qui réussit mieux l'assignation du genre grammatical. Du moins, les résultats ne sont absolument pas significatifs.

3.6 Les non-mots

Nous rappelons que, pour le test de reconnaissance des noms (Test ON : Oui/Non), les non-mots auxquels on a répondu O ont valu un point en moins pour les participants concernés. Nous rappelons aussi que la distribution des terminaisons des non-mots

correspondait aux trois catégories de Lyster (2006) selon que la terminaison était typiquement féminine (5 non-mots), masculine (5 non-mots) ou ambiguë (5 non-mots).

Pour ce qui est du test d'assignation du genre aux noms (Test MF : Masculin/Féminin), nous avons observé quelques résultats d'assignation du genre grammatical aux non-mots afin d'appuyer les résultats d'assignation du genre grammatical aux vrais noms. Comme ce sont des non-mots, nous n'avons évidemment pas pu observer quoi que ce soit concernant l'interprétation de résultats en faveur d'une explication d'assignation du genre par le transfert de la L1. Nous nous sommes donc penché sur la stratégie de l'assignation du genre par les terminaisons. Il faut admettre ici que, puisque nous n'avons pas tenu compte d'une potentielle congénéarité des terminaisons (vrais mots ou non-mots) avec les terminaisons espagnoles, il nous manque certains éléments d'interprétation des résultats. Cependant, nous rappelons que nous nous sommes assuré que les non-mots ne ressemblaient pas particulièrement à des noms espagnols, et ce, avec l'aide de deux hispanophones du département de linguistique de l'UQAM.

Tableau 3.11

Pourcentage de réussite au test d'assignation du genre, selon les terminaisons en -e muet final, pour les dix non-mots à terminaisons marquées pour le genre

	Ne (5)		e (5)		Total (10)	
	moyenne	écart type	moyenne	écart type	moyenne	écart type
H	88,6	18,4	62,1	28,5	75,4	16,4
S	83,1	16,9	64,1	16,0	73,6	13,8

Comme les résultats les plus significatifs étaient ceux qui concernaient les terminaisons en -e muet final, c'est ce que nous avons observé sur les non-mots. Une des choses intéressantes à considérer avec les non-mots est le fait que les résultats des hispanophones, comme ceux des sinophones, n'étaient influencés ni par l'effet du transfert ni par l'effet de la fréquence des mots ou de la connaissance des participants de ces items.

Dans le Tableau 3.11, on n'observe aucune différence significative entre les groupes linguistiques en ce qui concerne l'assignation préférée par chacun des participants aux dix non-mots ayant une terminaison marquée pour le genre, que cette terminaison ait ou non un *-e* muet final (et que la terminaison en question soit typiquement masculine ou typiquement féminine). Cependant, on peut considérer une différence significative entre les résultats de présence et d'absence de *-e* muet final (chez les hispanophones, $p = 0,0004$; et chez les sinophones, $p < 0,0001$).

Cela signifie donc que, peu importe le groupe linguistique, les participants ont eu de meilleurs résultats lorsque les non-mots ne possédaient pas de *-e* muet que lorsqu'ils en possédaient, quelle que soit la nature de la terminaison (typiquement masculine ou typiquement féminine). Ce constat accentue l'idée que le genre grammatical des noms sans terminaison en *-e* muet (Ne) serait plus facile à attribuer que pour les noms pourvus d'une terminaison en *-e* muet (e). Cela confirme les observations de Lyster (2006) à ce sujet. De plus, le manque de différence significative entre les deux groupes linguistiques pour les résultats globaux montre bien que les hispanophones, lorsqu'ils n'ont pas accès au transfert de leur LI, se comportent de façon similaire aux sinophones pour assigner le genre grammatical, c'est-à-dire qu'ils assignent ce dernier en utilisant la stratégie du recours à la terminaison en *-e* ou non.

Pour ce qui est des cinq non-mots qui n'ont pas de terminaison marquée pour le genre (c'est-à-dire qu'ils ont une terminaison ambiguë), mais qui peuvent se terminer ou non par un *-e*, les résultats vont sensiblement dans la même direction. Lorsque les non-mots ont une terminaison en *-e* muet final, les deux groupes linguistiques lui attribuent plutôt le genre féminin (bauxique = 42/57, néramme = 48/60, vérode = 47/60). Lorsque les non-mots ne possèdent pas une terminaison en *-e* muet final, les deux groupes linguistiques lui attribuent plutôt le genre masculin (mâchoix = 46/60, fison = 49/60).

Globalement, l'observation des résultats d'assignation du genre aux non-mots nous permet d'appuyer davantage et de confirmer nos hypothèses en ce qui concerne l'influence prédominante de la présence ou de l'absence des terminaisons en *-e* muet final (que ce *-e*

fasse partie d'une terminaison typiquement masculine ou typiquement féminine) comme stratégie d'assignation du genre grammatical en français.

3.7 Les non-réponses

Quelques items (noms ou non-mots) n'ont pas été pris en considération dans les résultats parce qu'il n'y a pas eu de réponse, pour ces items, à l'intérieur des 10 000 millisecondes allouées.

Pour le test de reconnaissance, 10 items, pour tous les participants et tous les mots, font partie de cette catégorie. Les non-mots auxquels les sujets n'ont pas répondu sont les suivants : *lipanage*, *caricité* et *cramance*. Les noms auxquels les sujets n'ont pas répondu sont les suivants : *blasphème*, *choix*, *céréale*, *montre*, *couleur*, *rançon* et *lâcheté*. Nous avons noté seulement deux non-réponses pour les hispanophones et huit pour les sinophones.

Pour le test d'assignation du genre, 11 items, pour tous les participants et tous les mots, font partie de cette catégorie. Les non-mots auxquels les sujets n'ont pas répondu sont les suivants : *néméade* et *bauxique* (3 fois). Les noms auxquels les sujets n'ont pas répondu sont les suivants : *tracteur*, *clarinette*, *gorgée*, *rançon*, *joue* et *nageoire* (2 fois). Nous avons noté cinq non-réponses pour les hispanophones et six pour les sinophones.

Il est intéressant de constater que sur les onze non-réponses, neuf se terminent par un -e muet final et seulement deux, *tracteur* et *rançon*, n'en possèdent pas. Cela appuie l'idée que, lorsqu'un nom possède un -e final muet, il est moins facile d'assigner le genre grammatical et les participants ont donc eu plus de difficulté à fournir une réponse quelconque (voir Lyster, 2006, au chapitre I).

Autant pour le test de reconnaissance que pour le test d'assignation du genre, les noms auxquels les sujets n'ont pas répondu ont des fréquences majoritairement sous la moyenne des fréquences dans Baudot. En effet, sur un total de 14 occurrences de noms auxquels les sujets n'ont pas répondu (excluant les non-mots), seulement deux (*choix*, 152 et *couleur*, 254) ont des fréquences au-dessus de la moyenne des fréquences de Baudot des noms de notre corpus (66,3). Cela peut être une piste d'explication de l'hésitation des participants.

3.8 Les résultats marginaux

Nous avons observé les quelques résultats marginaux que nous avons trouvés. Par « résultats marginaux », nous entendons ceux des participants, pour un test donné, qui dépassent l'écart de la majorité des autres participants (au-dessus ou au-dessous). Nous avons analysé ces résultats autant chez les hispanophones que chez les sinophones. En nous référant aux données sociobiographiques des participants concernés, nous n'avons pas pu trouver de tendance explicative. Certains participants possédaient certaines connaissances d'autres langues, notamment des langues qui détenaient des genres grammaticaux. Parmi les participants aux résultats marginaux, nous avons repéré ce cas de figure autant dans les résultats au-dessus qu'au-dessous de la moyenne. Il est donc impossible d'analyser cet aspect.

Il est à noter, de façon anecdotique, que deux participants sinophones ont eu des résultats relativement faibles, mais que ces résultats s'expliquent assez bien : un était enrhumé le jour du test et n'était donc pas à son meilleur et l'autre comprenait à peine le français ou l'anglais (le même participant qui a obtenu les résultats globaux les plus bas aux deux tests), à l'oral. Pour ce qui est des quelques participants aux résultats nettement au-dessus de la moyenne, nous avons constaté qu'ils semblaient assez souvent très avancés en français (hispanophones et sinophones), du moins à l'oral. En dehors de ces cas de figure exceptionnels, il est impossible de nous avancer davantage dans des explications qui seraient plutôt spéculatives. De toute manière, ces résultats marginaux sont, comme le terme l'indique, exceptionnels.

3.9 Conclusion

Pour conclure, il nous semble que les résultats les plus significatifs sont les suivants :

Lorsque les langues sont similaires (ici, espagnol et français), les résultats nous font pencher pour l'hypothèse de la stratégie du recours au transfert de la L1 comme première stratégie utilisée pour assigner le genre grammatical. Lorsque les langues ne sont pas similaires (ici, chinois et français), les résultats nous font plutôt pencher pour l'hypothèse de la stratégie du recours aux terminaisons en *-e*. De plus, cette dernière stratégie semble succéder à la première, pour les hispanophones, lorsque le recours au transfert n'est plus

possible, comme pour les non-mots. L'assignation du genre selon que les terminaisons des noms sont marquées ou non pour le genre a son importance, et ce, pour les deux groupes linguistiques. Finalement, la fréquence des noms (et le fait de connaître ou non le nom) aurait aussi un impact.

CHAPITRE IV

DISCUSSION GÉNÉRALE

4.1 Introduction

Dans ce dernier chapitre, nous retournerons à une perspective plus large. Premièrement, nous discuterons les liens entre les deux principales études sur lesquelles s'appuie notre recherche : Sabourin, Stowe et De Haan (2006) et Lyster (2006). Ensuite, nous discuterons d'autres observations ponctuelles. Finalement, nous parlerons des limites de notre recherche et des recommandations que nous avons à faire pour les futures recherches sur le sujet.

4.2 Les liens avec les deux études importantes

Rappelons que nous nous sommes basé principalement sur deux études. Nous avons, tout d'abord, jugé pertinente l'étude de Sabourin, Stowe et De Haan (2006), pour ce qui est de la question du transfert qui opérerait dans l'assignation du genre grammatical. Pour ce qui est de l'effet des terminaisons sur l'assignation du genre, l'étude de Lyster (2006) était celle qui nous a semblé la plus appropriée. Mentionnons, ici qu'en dépit du fait que ce sont les études de Tucker, Lambert et Rigault (1977) qui sont parmi les premières à développer l'idée du lien entre les terminaisons et le genre grammatical en français, nous nous sommes limité à celle de Lyster (2006) parce que les études de Tucker, Lambert et Rigault (1977) visaient surtout les locuteurs natifs du français alors que l'étude de Lyster se situe dans le contexte de l'apprentissage du français langue seconde.

4.2.1 L'effet du transfert sur l'assignation du genre

Rappelons au lecteur que l'étude de Sabourin, Stowe et De Haan (2006) compare les résultats d'assignation du genre pour trois groupes d'apprenants du néerlandais : des locuteurs de l'allemand (25 participants), des locuteurs de trois langues romanes : le français,

l'espagnol et l'italien (21 participants) et des locuteurs de l'anglais (24 participants). Rappelons aussi que les chercheurs de cette étude voulaient, entre autres, vérifier si le transfert profond (posséder le concept de genre dans la langue source) donnait de meilleurs résultats que le transfert superficiel (non seulement posséder le concept de genre dans la langue source, mais aussi avoir les mêmes genres que dans la langue cible (avec une distribution parfois similaire) et avoir de nombreuses similarités avec la langue cible (les congénères)).

Dans notre recherche, les hispanophones correspondent au groupe de la langue source, le français étant la langue cible. Le potentiel de transfert entre ces deux langues correspond à un transfert à la fois profond et superficiel, à savoir que l'espagnol possède le concept de genre, connaît les mêmes deux genres qu'en français et a beaucoup de similarités avec le français, notamment au niveau du lexique (les congénères). La distinction entre transfert profond et transfert superficiel n'est donc pas, comme nous l'avons déjà dit, observée dans notre recherche.

Dans notre recherche, nous confirmons tout de même les résultats de Sabourin, Stowe et De Haan (2006) pour le néerlandais, à savoir que les résultats suggèrent fortement un recours au transfert de la L1 lors de l'assignation du genre grammatical (les résultats des hispanophones). En effet, les résultats de l'étude de Sabourin, Stowe et De Haan (2006) allaient dans la même direction, puisque les résultats d'assignation du genre en néerlandais étaient significativement supérieurs pour les participants des groupes linguistiques possédant un genre (les locuteurs de l'allemand et des langues romanes), et inférieurs pour le groupe des anglophones, l'anglais ne possédant pas de genre.

Dans notre recherche, il est clair que les hispanophones se fient au genre de leur langue, surtout lorsque les noms observés sont congénères avec ceux de leur L1. Cependant, contrairement aux résultats d'assignation des anglophones de la recherche de Sabourin, Stowe et De Haan (2006), dans notre recherche, les sinophones n'obtiennent pas des résultats globaux inférieurs aux hispanophones bien qu'ils n'aient pas accès à la possibilité du transfert

de leur L1. C'est que la question de l'assignation du genre selon les terminaisons des noms entre en jeu pour le français langue seconde, comme nous l'avons vu⁴.

Pour ce qui est de l'effet de la fréquence des noms sur l'assignation du genre, nous avons aussi, semble-t-il, des résultats similaires à ceux de l'étude de Sabourin, Stowe et De Haan (2006). Dans notre recherche, l'hypothèse selon laquelle la haute fréquence des noms est corrélée à la bonne assignation du genre est confirmée. Rappelons tout de même que, dans la recherche de Sabourin, Stowe et De Haan (2006), tous les participants avaient été testés au préalable afin d'être assurément des apprenants avancés en néerlandais. Dans notre recherche, nous n'avons pas prédéterminé le niveau en français des participants, en autant qu'ils nous comprenaient lorsque nous leur parlions. Cela dit, il en est resté que nos participants étaient tous des apprenants intermédiaires ou avancés en français.

Finalement, l'hypothèse de l'assignation du genre non marqué par défaut est infirmée dans notre recherche, tandis qu'elle est confirmée dans l'étude de Sabourin, Stowe et De Haan (2006). En effet, nous avons vu que les erreurs des apprenants (hispanophones ou sinophones) ne tendaient pas vers une assignation du masculin (genre non marqué) par défaut. Contrairement aux résultats de notre recherche, le déterminant correspondant au genre commun (« de »), le genre non marqué en néerlandais, est assigné par défaut lorsque les noms présentés sont de moyenne fréquence d'occurrence.

En somme, n'oublions pas qu'une comparaison entre une recherche sur l'assignation du genre grammatical en néerlandais et une autre sur cette même assignation, mais en français, a ses limites. Néanmoins, il est intéressant de constater des résultats similaires pour ce qui est d'un des aspects majeurs de nos recherches respectives : l'effet du transfert sur l'assignation du genre grammatical.

⁴ Nous ne savons pas si la question des terminaisons des noms est pertinente pour le néerlandais, mais il nous semble que les résultats des anglophones, dans la recherche de Sabourin, Stowe et De Haan (2006), n'ait pas été influencés positivement dans cette direction.

4.2.2 L'effet des terminaisons sur l'assignation du genre

Comme l'article de Sabourin, Stowe et De Haan (2006) nous a servi de base pour discuter l'effet du transfert de la L1 lors de l'assignation du genre grammatical, l'article de Lyster (2006) nous a servi de base pour discuter l'effet des terminaisons sur cette même assignation.

Rappelons au lecteur que l'article de Lyster (2006) était une étude statistique des représentations orthographiques des rimes des noms du *Robert Junior Illustré* afin de repérer des associations entre les terminaisons des noms et le genre grammatical de ces derniers. Cette étude se voulait génératrice de patrons lexico-grammaticaux pour fins pédagogiques (en français langue seconde). Les résultats révélaient que 81 % de tous les noms féminins et 80 % de tous les noms masculins du corpus possédaient une terminaison qui prédisait fortement le genre. L'auteur classifiait les 239 terminaisons valables pédagogiquement selon qu'elles étaient typiquement féminines, typiquement masculines ou typiquement ambiguës. Les terminaisons, pour chacun des groupes, devaient représenter 90 % et plus du corpus pour être classées comme typiquement féminines ou masculines. Celles qui n'obtenaient pas ce résultat faisaient partie des terminaisons typiquement ambiguës.

Dans notre recherche, nous avons confirmé notre hypothèse concernant l'effet des terminaisons marquées pour le genre sur la bonne assignation de celui-ci. Cela va donc dans la même direction que les données linguistiques de Lyster (2006), à savoir que les résultats démontrent, de manière empirique, que les terminaisons telles que classifiées par Lyster (2006) incitent effectivement les apprenants à assigner le genre approprié.

Nous avons aussi enregistré des résultats spectaculaires chez les sinophones concernant l'assignation du genre selon que la terminaison des noms se termine en *-e* muet final ou non. Lyster (2006) observe que les terminaisons finissant par un *-e* muet sont majoritairement féminines et que les terminaisons ne finissant pas par un *-e* muet sont presque qu'exclusivement masculines. Cependant, nous avons vu que beaucoup de terminaisons ambiguës possédaient également une terminaison en *-e* muet. Nos résultats tendent donc à s'aligner avec les données linguistiques considérées par Lyster (2006).

Finalement, comme pour l'étude de Sabourin, Stowe et De Haan (2006), Lyster (2006) dit avoir déjà observé que le genre non marqué du français, le masculin, avait tendance à être le genre majoritaire des résultats d'erreurs des apprenants. Encore une fois, nous rappelons que nous avons obtenu, ici, des résultats contraires. Nous ne pouvons pas, d'après nos résultats, confirmer cette hypothèse de l'assignation du genre masculin par défaut, comme pour Sabourin, Stowe et De Haan (2006) et Lyster (2006).

4.3 Une observation

Une des pistes d'interprétation des résultats qui nous est apparue est que les sinophones se fient beaucoup plus à l'orthographe, stratégie visuelle, qui découle en partie de leur façon d'approcher leur propre langue et qui est liée au type d'apprentissage scolaire. En effet, en chinois, on doit apprendre par cœur une grande quantité d'idéogrammes. De plus, les Chinois valorisent l'écrit dans leurs traditions. Cela les a amenés à recourir à une stratégie de l'écrit. Bien sûr, comme il ne possédaient pas vraiment d'indices dans leur langue pour les aider dans leur reconnaissance des noms français ou dans leur assignation du genre grammatical, il devaient obligatoirement se pencher sur l'étude de la langue à apprendre, c'est-à-dire le français. De plus, les tests que nous avons fait passer étaient sur écran d'ordinateur. Les mots étaient écrits, donc lus, et non entendus. Cela est vrai pour les deux groupes linguistiques, mais les hispanophones se sont fiés probablement beaucoup plus au sens et à la parenté linguistique qu'aux marques écrites. Cette affirmation a, en effet, été appuyée par nos résultats.

4.4 Les limites et les recommandations

Nous allons maintenant brièvement faire une critique de notre recherche. Nous pensons que, de manière générale, nous avons bien mené cette recherche à terme. Par contre, il reste toujours des angles morts à considérer lors de recherches ultérieures ou des aspects qui ne pouvaient pas être traités en même temps.

Nous pensons tout d'abord que nous aurions pu vérifier, lors de l'analyse des données, le genre des noms utilisés dans les différentes variantes dialectales de l'espagnol des sujets, afin qu'il n'y ait pas une distribution du genre différente dans un dialecte particulier. Cet aspect nous semblait dépasser de loin le cadre de notre mémoire.

Nous croyons aussi que les résultats dépendent, bien sûr, de plusieurs variables. Notamment, « [l]a compétence lexicale de l'apprenant dans sa L1 intervient immédiatement dans l'identification de congénères car, comment reconnaître un congénère en L2 quand sa contrepartie en L1 n'est pas connue par lui ? » (Theophanous, 2000, p.88). Nous ne pouvions pas vérifier les compétences en L1 de chacun de nos participants, bien que cela eut été plus fiable.

« La fréquence du mot de la L1 est un bon critère pour prévoir les chances qu'a sa contrepartie congénère en L2 d'être reconnue comme telle » (Theophanous, 2000, p. 100). En effet, la reconnaissance d'un congénère dépend certainement, encore une fois, de la fréquence, voire de la connaissance de ce nom dans la L1. Cela est une avenue qui pourrait être étudiée plus en profondeur.

Theophanous (2000) donne plusieurs facteurs pouvant influencer la reconnaissance des congénères, mais nous n'avons pas approfondi ces avenues lors de notre sélection des congénères.

Même si « [o]n pourrait s'attendre à ce que l'utilisation des deux articles singuliers indéfinis UN et UNE, comparativement aux étiquettes MASCULIN et FÉMININ, donne lieu à des latences de réponses significativement plus courtes » (Desrochers, Paivio et Desrochers, 1989, p. 65), nous croyons que les étiquettes M et F étaient convenables. En effet, dans notre recherche, nous n'avons pas pris en compte le temps de réponse comme pour la recherche de Desrochers, Paivio et Desrochers (1989). Nous croyons, en effet, que les articles « un » et « une » auraient empiété sur le terrain de l'accord en genre. De plus, les articles sont parfois confondus, du point de vue de la prononciation, ce qui aurait pu fausser les résultats. Même si « [c]es étiquettes sont habituellement introduites à l'école et utilisés dans un discours réflexif sur la grammaire [...] » (Desrochers, Paivio et Desrochers, 1989, p. 65), et que nous ne sommes pas à l'abri d'une incompréhension ou d'une confusion de l'apprenant, nous maintenons notre pensée qu'elles sont plus fiables.

Dans toute recherche, il y a plusieurs facteurs qui échappent au chercheur. Parmi ces autres facteurs, il y a tous les facteurs extra-linguistiques comme la motivation, le contexte d'apprentissage (classe, milieu naturel, grand groupe, petit groupe, individuel, etc.), etc. Ces

facteurs non structurels semblent opérer indépendamment du transfert (Odlin, 1989, p. 129). Cependant, nous croyons qu'ils ont toujours un certain impact sur des résultats de recherche. De plus, il semble y avoir d'autres facteurs non structurels qui interagissent avec le transfert, ceux en relation avec la variation individuelle : l'âge de l'apprenant, sa personnalité, son aptitude pour la « phonetic mimicry », sa compétence en langues, son niveau de littératie, etc. De plus, la conscience linguistique et le contexte social entourant l'apprenant sont d'autres facteurs importants (Odlin, 1989, p. 129-150).

Comme nous l'avons mentionné précédemment, nous aurions pu tenir compte, de façon plus raffinée, du détail des terminaisons des noms espagnol. En effet, nous nous sommes concentré sur les terminaisons en français, mais les noms espagnols possèdent aussi des terminaisons. Cela reste à être exploré dans des recherches ultérieures.

De plus, nous avons choisi d'exclure plusieurs catégories de noms (les noms à initiale vocalique, par exemple, ou les noms animés). Il serait fort intéressant de tester ces autres catégories de noms, mais dans des recherches séparées.

Comme nous l'avons déjà mentionné dans le chapitre 3, nous avons vérifié si des résultats marginaux n'étaient pas causés par l'effet d'un transfert potentiel d'autres langues apprises par l'apprenant. « Transfer is the influence resulting from the similarities and the differences between the target language and **any other language** that has been previously (and perhaps imperfectly) acquired » (Odlin, 1989, p.27). Nous aurions pu pousser les investigations concernant le transfert d'une troisième langue chez le même apprenant.

Nous nous sommes concentré, dans ce travail, sur l'assignation du genre, mais tout un pan de recherche reste à faire sur l'accord en genre. De plus, les noms choisis étaient écrits, mais nous aurions pu vérifier presque les mêmes hypothèses, et avec les mêmes noms, mais à l'oral (enregistrement sonore).

Finalement, nous devons ajouter que le choix de nos groupes linguistiques dépendait en partie de notre possibilité de trouver suffisamment de participants pour que la recherche soit valable. Bien sûr, nous pourrions recommencer la recherche avec d'autres groupes linguistiques.

4.5 Conclusion

Dans ce chapitre, nous avons comparé les résultats de notre recherche à ceux des deux études principales sur lesquelles nous nous sommes appuyé, Sabourin, Stowe et De Haan (2006) et Lyster (2006). Nous avons aussi parlé d'une observation ponctuelle que nous avons faite concernant nos résultats. Ensuite, nous avons décrit les limites auxquelles nous avons été confronté tout au long de notre recherche et nous avons ouvert nos réflexions à des aspects à explorer plus amplement, ou sur des projets à venir.

CHAPITRE V

IMPLICATIONS PÉDAGOGIQUES

Nous allons maintenant discuter les différents aspects concernant les implications, les retombées pédagogiques à la lumière de nos lectures et des résultats obtenus dans notre recherche.

5.1 Notre modèle général d'acquisition du genre grammatical

Nous allons faire ici une ébauche de théorie d'acquisition du genre grammatical d'après nos résultats. Nous pourrions extrapoler l'hypothèse que ce modèle serait applicable à d'autres éléments langagiers que le genre grammatical. Considérant l'acquisition d'une langue seconde, nous sommes toujours en présence de deux langues (du moins pour commencer). Il y a toujours la langue source, celle de l'apprenant, et la langue cible, celle à apprendre. Il y a possibilité que l'aspect à acquérir (en l'occurrence, le genre grammatical) soit plus ou moins similaire de la langue source à la langue cible. Plus cet élément langagier est similaire dans les deux langues, plus il y aura potentiellement recours au transfert de la langue source. Inversement, moins cet élément langagier est similaire dans les deux langues, moins il y aura de recours au transfert, et plus l'apprenant recourra à une stratégie qui tient compte de la langue cible.

L'exemple de notre recherche est extrême puisque les deux groupes linguistiques choisis sont à l'opposé en ce qui concerne la similarité du genre grammatical comme élément linguistique. En espagnol, les genres sont les mêmes qu'en français (féminin et masculin) et ont une distribution très similaire à celle du français, tandis qu'en chinois, il n'y a carrément pas de genre. Dans le cas où l'apprenant n'aurait pas accès à sa langue pour déterminer l'assignation du genre des noms de la langue cible (ici, les sinophones), il développerait une stratégie tenant plutôt compte de la langue cible (ici, l'assignation du genre selon les

terminaisons en *-e* muet final). Cependant, la langue cible aurait toujours son impact, et ce, peu importe la langue source (exemple de l'impact des terminaisons marquées pour le genre sur les résultats des hispanophones et des sinophones).

Dans ce modèle, la fréquence doit être prise en compte. Plus un mot est fréquent, plus l'apprenant lui assignera le bon genre grammatical. De plus, dans le cas des apprenants ayant recours au transfert de leur langue, la reconnaissance des noms aurait une importance et accentuerait la bonne assignation du genre grammatical. En effet, la reconnaissance des noms (comme la congénérarité de ceux-ci avec la langue source) était un facteur clé pour les hispanophones, mais pas pour les sinophones.

Tentons maintenant de traduire ce modèle en applications pédagogiques.

5.2 Les considérations générales

Comme le dit Odlin (1989, p. 157) « [...] cross-linguistic influence has considerable potential to affect the course of second language acquisition both inside and outside the classroom. » Cependant, ce ne sont pas toutes les langues sources et tous les aspects de la langue qui peuvent être transférés. Comme nous l'avons vu dans les résultats de notre recherche, pour ce qui est du genre grammatical, les sinophones ne peuvent transférer absolument rien. Qu'en est-il, alors des stratégies que l'enseignant peut appliquer dans la salle de classe, et ce, à la lumière de nos résultats?

D'abord, nous pouvons nous poser la question suivante : dans quel sous-domaine de l'apprentissage ou de l'enseignement nos résultats peuvent-ils servir? L'assignation du genre concerne moins la compréhension orale ou la lecture, mais plutôt tout ce qui a trait à la production, orale ou écrite. Bien sûr, nous croyons que les champs d'application de nos résultats sont multiples. Il nous faudrait un mémoire complet pour discuter chaque aspect, ne serait-ce que la différence entre la production orale et la production écrite.

De plus, deux questions s'imposent à nous : quelles sont les implications possibles dans un contexte d'apprentissage dit « naturel » et quelles sont les implications pédagogiques possibles du point de vue de l'enseignement en milieu formel? Nous pensons qu'il y a un continuum de possibilités entre les deux contextes d'apprentissage et qu'il nous est

impossible de discuter chaque aspect dans le contexte de ce mémoire. Nous allons donc plutôt parler du contexte de l'enseignement formel, avec un enseignant.

Il nous a semblé que les spécialistes de la didactique du vocabulaire en français langue seconde ne parlaient presque jamais de la façon d'enseigner le genre grammatical. « Since vocabulary plays a crucial role in all uses of language, teachers should take a special care to monitor what words individual students may find easy or difficult in reading materials or in spoken language in the classroom » (Odlin, 1989, p. 160). Il est difficile de savoir si les professeurs devraient enseigner le genre grammatical de façon explicite et formelle, surtout quand vient la question de l'assignation du genre grammatical. Nous voyons plus probable cet enseignement formel pour ce qui est de l'accord en genre. « [...] [t]eachers must be concerned not only with forms and functions but also with the learning process. Although transfer is only one aspect of that process, it is a crucial one » (Odlin, 1989, p. 161). Nous pensons, donc, que l'acquisition du genre grammatical en français est une question de processus, pas seulement une question formelle. Cependant, comme nous l'avons vu, il n'y a pas seulement le transfert de la L1 qui est un processus potentiel, mais aussi certaines stratégies reliées aux terminaisons des noms, notamment, des *-e* muet finaux. Nous croyons tout de même que l'acquisition d'une langue seconde dépend plus d'un enseignement formel et explicite que l'acquisition de la langue première. Le genre grammatical n'échappe pas à cette règle, selon nous, bien qu'il y ait toujours des exceptions d'apprenants qui acquièrent presque tous les aspects d'une langue seconde sans enseignement explicite.

Penchons-nous, maintenant, sur la façon d'enseigner la bonne assignation du genre grammatical. « Fries, Lado and other contrastive theoreticians believed that Spanish-speaking students, for example, needed textbooks and other materials very different from those needed by Chinese-speaking students » (Odlin, 1989, p. 161). Comme les auteurs de la citation de Odlin (1989), nous pensons qu'il y a différentes façons d'enseigner pour différents groupes linguistiques. Bien sûr, il n'est pas toujours aisé, concrètement, de nous adapter à chaque groupe. De plus en plus, le professeur de langue seconde ou étrangère est confronté à des groupes multiethniques. Cependant, nous sommes d'accord avec Odlin (1989, p. 163) pour dire que « [I]t is always possible to learn the language of students or at least enough about a student's language to develop contrastive descriptions. » En effet, nous croyons,

contrairement à bien des enseignants de langues secondes (anecdotique), que, parfois, il est préférable, voire très aidant, de nous référer à la langue de l'apprenant. Par exemple, « [d]u point de vue pédagogique, devant l'immensité du vocabulaire, les congénères peuvent figurer parmi les premiers mots à être enseignés » (Theophanous, 2000, p.87). Cette dernière affirmation concerne l'enseignement du vocabulaire, mais peut-être pourrait-elle concerner aussi l'enseignement du genre grammatical des noms en français, par exemple, en enseignant d'abord les noms qui ont le même genre dans les deux langues.

5.3 Nos suggestions d'applications pédagogiques

Nous pensons donc que, pour l'enseignement du genre grammatical en français, il serait intéressant de considérer l'enseignement aux hispanophones de façon différente de l'enseignement aux sinophones. À la lumière des résultats de notre recherche, nous pourrions, par exemple, dire aux hispanophones de suivre leurs intuitions (souvent le genre de leur langue) tout en leur donnant une courte liste de noms assez fréquents qui sont congénères avec les noms espagnols, mais qui ont le genre inverse d'une langue à l'autre. Bien sûr, nous sommes conscient que la réalité de l'enseignement formel (répartition inégale des groupes linguistiques, groupes multiethniques, etc.) ne permet pas toujours ces conditions. De plus, nous croyons que l'enseignement en groupes hétérogènes a ses avantages (à d'autres niveaux). Nous ne disons pas ici que les groupes linguistiques devraient à tout prix être séparés, dans des groupes différents (si cela était possible).

Comme nous l'avons vu dans la section sur l'espagnol (section 1.7.1) du chapitre 1, selon Teschner (1987, p. 257), dans son étude comparative du genre des congénères entre l'espagnol et le français, 75,43 % des noms sont congénères. Parmi ceux-ci, 92,38 % ont le même genre. Seulement 7,62 % (sur les 746 noms listés dans Teschner) des congénères n'auraient pas le même genre qu'en français. Dans ce 7,62 % de noms congénères qui n'ont pas le même genre qu'en français, seulement une partie est probablement constituée de noms fréquents et utiles. De plus, il faut savoir que nos résultats, lors des tests statistiques, concernent une répartition équilibrée des noms selon qu'ils sont congénères ou non et qu'ils possèdent le même genre ou non dans les deux langues. De cette façon, avec les données de Teschner (1987) en tête, les hispanophones auront probablement une meilleure assignation du genre en français, de manière générale, confrontés au plein lexique français. Nous rappelons

au lecteur que les résultats globaux d'assignation du genre pour les hispanophones étaient de 62,7 % (écart type de 6,9) et de 60,2 % pour les sinophones (écart type de 5,3). En réalité, nous sommes d'opinion que, si notre liste de noms avait reflété les proportions décrites par Teschner (1987) (peu de congénères n'ayant pas le même genre qu'en français) les hispanophones auraient tout probablement eu de meilleurs résultats que pour les proportions lexicales de nos tests actuels (classement binaire).

De la même façon, nous pourrions enseigner aux sinophones de suivre ce qui semble être leur « tendance naturelle », c'est-à-dire d'assigner le genre féminin aux noms qui se terminent en *-e* muet final, et d'assigner le genre masculin aux noms qui ne se terminent pas en *-e* muet final. Il faudrait en même temps leur donner une liste d'exceptions à cette règle parmi des noms assez fréquents et utiles. De la même manière que pour les hispanophones, nous croyons que les résultats d'assignation du genre grammatical des sinophones seraient supérieurs face à l'ensemble du lexique que lors de nos tests, puisque la répartition des genres selon les terminaisons en *-e* ou non n'est pas autour de 50 %, comme dans notre corpus, mais supérieure (voir les tableaux de Lyster (2006), au chapitre 1).

Nous croyons que, de façon générale, il serait possible, surtout lorsque les enseignants du français langue seconde se retrouvent en face de groupes multilinguistiques, d'enseigner la règle des *-e* muet finaux, telle qu'elle est employée par les sinophones, mais en rajoutant les exceptions suivantes : *-ion*, *-té* et *-isme*, *-age*. En effet, dans les tableaux de Lyster (2006, section sur l'article de Lyster du chapitre 1), il semble que les terminaisons *-ion* et *-té* soient les deux seules terminaisons sans *-e* muet final parmi les terminaisons typiquement féminines. De plus, il semble que les terminaisons *-isme* et *-age* soient les deux terminaisons les plus fréquentes (sur sept terminaisons) avec un *-e* muet final parmi les terminaisons typiquement masculines. Bien sûr, cette règle de l'assignation selon les terminaisons en *-e* muet final, sauf pour *-ion*, *-té*, *-isme* et *-age* ne donnerait pas de résultats parfaits, surtout lorsque les noms se termineraient par un *-e* muet final (notamment parce qu'il resterait les terminaisons ambiguës, principalement en *-e* muet final), mais il nous semble qu'elle soit beaucoup plus valable que l'idée traditionnelle que le genre est arbitraire et ne s'enseigne pas. De plus, elle nous semble plus simple que les longues listes de terminaisons et d'exceptions de plus en plus fréquentes que l'on retrouve dans les manuels de français langue seconde, et

ce, même si les terminaisons sont significativement porteuses du genre, comme le démontre Lyster (2006).

Bien sûr, il faudrait penser à mieux former les professeurs de français langue seconde à la lumière des nouveaux résultats de la recherche dans le domaine. L'aspect culturel est aussi une variable que l'on oublie souvent. Peut-être que certaines approches de l'enseignement du genre fonctionneraient avec certains groupes linguistiques, mais pas avec d'autres. Nous croyons, de toute façon, que le genre grammatical est un aspect de la langue française qui est, plus souvent qu'autrement, escamoté dans l'enseignement, et qu'il y aurait un bon travail d'amélioration à obtenir à ce sujet.

5.4 Conclusion

Dans ce dernier chapitre, nous avons vu quelques pistes d'applications pédagogiques concernant le genre grammatical. Nous avons suggéré, notamment, qu'il serait beaucoup plus simple d'enseigner le genre d'une façon différente aux groupes d'hispanophones et aux groupes de sinophones. Cet enseignement tiendrait compte des différentes stratégies qui nous ont semblé utilisées dans notre recherche par chacun des groupes linguistiques.

CONCLUSION

Nous concluons en rappelant que notre recherche portait principalement sur l'effet du transfert lors de l'assignation du genre grammatical en français langue seconde, ainsi que sur une comparaison de cette stratégie (processus) potentielle avec celle de l'induction du genre par la terminaison des noms. Voilà pourquoi nous avons fait passer des tests à des hispanophones (hypothèse du recours au transfert potentiel) et à des sinophones (groupe contrôle pour ce qui est du recours au transfert).

Le choix des mots, classés de façon binaire selon quatre catégories (mêmes genres ou non qu'en espagnol, congénères ou non avec les noms espagnols, terminaisons marquées ou non pour le genre et terminaisons en *-e* muet final ou non) nous a permis de répondre à nos questions de recherche concernant les deux grands axes de celle-ci : l'effet du transfert et celui des terminaisons des noms (terminaisons marquées pour le genre et terminaisons en *-e* muet final). Nous avons pu répondre à d'autres questions de recherche, notamment celles concernant l'influence de la fréquence des noms, du niveau de connaissance du vocabulaire des participants et du sexe des apprenants sur l'assignation correcte du genre grammatical. De plus, nous avons tenté de vérifier s'il y avait une assignation du masculin (genre non marqué) par défaut, lors des erreurs des apprenants.

Cette recherche se compare à une partie de celle de Sabourin, Stowe et de Haan (2006), mais sur le français au lieu du néerlandais, pour ce qui est de l'effet du transfert sur l'assignation du genre. En ce qui concerne les terminaisons, nous nous sommes appuyés surtout sur l'article de Lyster (2006).

Finalement, même si les résultats totaux d'assignation du genre grammatical étaient sensiblement les mêmes pour les deux groupes linguistiques (les hispanophones et les sinophones), ces résultats ont été répartis fort différemment. Les résultats les plus significatifs

tendent à s'interpréter de la façon suivante : les hispanophones utilisent le genre de leur langue, et ce, surtout pour les noms congénères avec leur L1, ce qui leur donne d'excellents résultats d'assignation du genre lorsqu'ils possèdent le même genre en espagnol, mais qui les induit en erreur lorsqu'ils n'ont pas le même genre dans leur langue. Les sinophones, quant à eux, se basent sur une assignation du genre féminin lorsque la terminaison des noms français est en *-e* muet, tandis qu'ils assignent massivement le genre masculin lorsque les noms français ne possèdent pas de *-e* muet final. De plus, nous pensons que les résultats des deux groupes linguistiques seraient supérieurs à ceux des tests de notre recherche puisque l'ensemble statistique du lexique français donne raison aux apprenants (selon leur tendance d'assignation respective).

L'hypothèse concernant l'influence positive des terminaisons marquées pour le genre (typiquement féminines et typiquement masculines) sur la bonne assignation du genre grammaticale a été confirmée. L'hypothèse concernant l'influence de la fréquence des noms a aussi été confirmée, mais celle sur la connaissance en vocabulaire des apprenants n'a été confirmée que pour le groupe des hispanophones, vu la grande quantité de noms congénères entre l'espagnol et le français. Les hypothèses concernant le sexe des apprenants et celle qui a trait au masculin par défaut ont toutes deux été infirmées.

APPENDICE A

DOCUMENTS UTILISÉS

A.1	Formulaire de consentement	93
A.2	Formulaire de renseignements généraux	94
A.3	Consignes	95
A.4	Touches	98

A.1 Formulaire de consentement

J'accepte de me porter volontaire pour cette étude qui s'insère dans le cadre du mémoire de maîtrise de Frédéric Gareau. Cette recherche permettra de mieux comprendre l'acquisition du vocabulaire chez les apprenants adultes du français langue seconde.

Je comprends que je pourrai me retirer de l'étude à tout moment pour des raisons personnelles.

Je suis conscient que les renseignements contenus dans mon dossier de recherche (réponses aux tests, renseignements généraux) pourront être utilisés par le responsable de l'étude. Cependant, les éléments qui pourraient être confidentiels ne seront pas divulgués de manière publique d'une façon qui permettrait de m'identifier.

Votre nom :

Signature du participant :

Date :

A.2 Formulaire de renseignements généraux

Prénom : _____

Nom de famille : _____

Sexe : Homme ☐ Femme ☐

Âge : _____ ans

Première langue apprise : _____

Utilisez-vous encore cette langue? Oui ☐ Non ☐

Langue(s) d'usage : _____

Connaissances d'autres langues : _____

Pays de naissance : _____

Pays où vous avez passé la plus

grande partie de votre vie : _____

Consignes générales

Bonjour et merci de participer à ma recherche!

La durée totale est d'environ 45 minutes.

Répondez à tous les items.
(maximum : 10 secondes par item)

Ne vous arrêtez pas entre deux items.
(attendez aux pauses prévues)

Donnez toujours vos réponses
avec la même main.

Test 1

Connaissez-vous le sens des mots affichés ?

Si vous savez ce que le mot veut dire,
appuyez sur ◀ (Oui).

Si vous ne savez pas ce que le mot veut dire,
appuyez sur ▶ (Non).

Test 2

Connaissez-vous le genre des mots affichés ?

Si vous croyez que le mot est masculin
(ex. : le citron, un citron, mon citron, ce
citron), appuyez sur ◀ (M).

Si vous croyez que le mot est féminin
(ex. : la tarte, une tarte, ma tarte, cette tarte),
appuyez sur ▶ (F).

Répondez à chaque question,
même si vous n'êtes pas sûr de votre réponse
et même si vous ne connaissez pas le mot.

A.4 Touches



Oui



Non



Masculin

Féminin

RÉFÉRENCES

- Adjémian, Christian. 1983. « The transferability of lexical properties ». In *Language transfer in language learning*, sous la dir. de Susan Gass et Larry Selinker, p. 250-268. Rowley (MA) : Newbury House.
- Ard, Josh, et Taco Homburg. 1983. « Verification of language transfer ». In *Language transfer in language learning*, sous la dir. de Susan Gass et Larry Selinker, p. 157-176. Rowley (MA) : Newbury House.
- Barbaud, Philippe, Christian Ducharme et Daniel Valois. 1982. « D'un usage particulier du genre en canadien-français : la féminisation des noms à initiale vocalique ». *La Revue canadienne de linguistique*, vol. 27, no 2 (automne), p. 103-133.
- Baudot, Jean. 1992. *Fréquence d'utilisation des mots en français écrit contemporain*. Montréal : Les Presses de l'Université de Montréal, 431 p.
- Bloomfield, Leonard. 1933. *Language*. New York : Holt, Rinehart and Winston, 564 p.
- Corbett, Greville G. 1991. *Gender*. Cambridge : Cambridge University Press, 363 p.
- Desrocher, Alain, Allan Paivio et Sylvie Desrochers. 1989. « L'effet de la fréquence d'usage des noms inanimés et de la valeur prédictive de leur terminaison sur l'identification du genre grammatical ». *Revue canadienne de psychologie*, vol. 43, no 1, p. 62-73.
- Dubois, Jean, Mathée Giacomo, Louis Guespin, Christiane Marcellesi, Jean-Baptiste Marcellesi et Jean-Pierre Mével. « Marque ». In *Dictionnaire de linguistique*, éd. 2001.
- Dulay, Heidi, et Marina Burt. 1974. « Natural sequences in child second language acquisition ». *Language Learning*, vol. 24, p. 37-53.
- Ellis, Rod. 1994. « Language transfer ». Chap. in *The study of second language acquisition*, p. 299-345. Oxford : Oxford University Press.
- Franceschina, Florencia. 2005. *Fossilized second language grammars : The acquisition of grammatical gender*. Amsterdam/Philadelphia : John Benjamins, 288 p.
- Fries, Charles. 1949. « The Chicago investigation ». *Language Learning*, vol. 2, p. 89-99.

- Granger, Sylviane. 1993. « Cognates : An aid or a barrier to successful L2 vocabulary development ? ». *ITL : Review of Applied Linguistics*, no 99-100, p. 43-56.
- Greenberg, Joseph Harold. 1972. « Numeral classifiers and substantival number : Problems in the genesis of a linguistic type ». *Working Papers on Language Universals*, vol. 9, p. 1-39.
- Harris, James Wesley. 1991. « The exponence of gender in Spanish ». *Linguistic Inquiry*, vol. 22, no 1 (hiver), p. 27-62.
- Hellinger, Marlis, et Hadumod Bussmann (dir. publ.). 2001-2003. *Gender across languages*. 3 t. Amsterdam/Philadelphia : John Benjamins.
- Hockett, Charles Francis. 1958. *A course in modern linguistics*. New York : Macmillan, 621 p.
- Holmes, Virginia Margaret, et Bernadette Dejean de la Bâtie. 1999. « Assignment of grammatical gender by native speakers and foreign learners of French ». *Applied Psycholinguistics*, vol. 20, p. 479-506.
- Hudson, Grover. 2000. *Essential introductory linguistics*. Malden (MA) : Blackwell publishing, 533 p.
- Ibrahim, Muhammad Hasan. 1973. *Grammatical gender: Its origin and development*. The Hague : Mouton, 113 p.
- Kellerman, Eric. 1979. « Transfer and non-transfer: Where are we now? ». *Studies in Second Language Acquisition*, vol.2, no 1, p. 37-57.
- Lado, Robert. 1957. *Linguistics across cultures*. Ann Arbor : University of Michigan Press, 141 p.
- Lyster, Roy. 2004a. « How to predict the grammatical gender of 8 000 French nouns ». Communication présentée à la *American Association for Applied Linguistics*, Portland, OR.
- Lyster, Roy. 2004b. « Differential effects of prompts and recasts in form-focused instruction ». *Studies in Second Language Acquisition*, vol. 26, p. 399-432.
- Lyster, Roy. 2006. « Predictability in French gender attribution: A corpus analysis ». *French Language Studies*, vol. 16, p. 69-92.
- Mel'čuk, Igor' Aleksandrovic. 1974. « Statistics and the relationship between the gender of French nouns and their endings ». In *Essays on lexical semantics I*, sous la dir. de Viktor Ju. Rozenevejg, p. 11-42. Stockholm : Skriptor. [L'original russe, 1958]

- Odlin, Terence. 1989. *Language transfer : Cross-linguistic influence in language learning*. Cambridge : Cambridge University Press, 210 p.
- Sabourin, Laura, Laurie A. Stowe et Ger J. de Haan. 2006. « Transfer effects in learning a second language grammatical gender system ». *Second Language Research*, vol. 22, no 1, p. 1-29.
- Surridge, Marie. 1995. *Le ou la?: The gender of French nouns*. Coll. « Modern Languages in Practice », no 1. Clevedon (R-U) : Multilingual Matters, 86 p.
- Teschner, Richard V. 1987. « The gender of French and Spanish noun cognates : Some statistics and a list ». *La Revue canadienne des langues vivantes*, vol. 43, no 2, p. 256-266.
- Theophanous, Olga. 2000. « Identification des congénères : quels facteurs sont en jeu ? Un état de la question ». *ITL: Review of Applied Linguistics*, vol. 131, no 132, p. 85-105.
- Tucker, Richard D., William E. Lambert et André A. Rigault. 1977. *The French speaker's skill with grammatical gender : An example of rule-governed behaviour*. The Hague : Mouton, 131 p.

OUTILS INFORMATIQUES

Bastien, Michel. *LÉA*
<http://microbe.ca/Lea.htm>

Microsoft Access

RAEi (Real Academia Española informatisée)
<http://www.rac.es/RAE/Noticias.nsf/Home?ReadForm>

TLFi (Trésor de la Langue Française informatisé)
<http://atilf.atilf.fr/tlfi3.htm>

Wikipedia (anglais)
http://en.wikipedia.org/wiki/Measurc_word#Japanese